# CHARACTERIZATION OF THE *MYCOBACTERIUM TUBERCULOSIS* ZINC METALLOPEPTIDASE ZMP1 AND GENERATION OF VACCINE STRAINS AGAINST TUBERCULOSIS

**Dissertation**
**zur**
**Erlangung der naturwissenschaftlichen Doktorwürde**
**(Dr. sc. nat.)**

**vorgelegt der**
**Mathematisch-naturwissenschaftlichen Fakultät**
**der**
**Universität Zürich**

**von**
**Agnese Petrera**

**aus**
**Italien**

**Promotionskomitee**

**Prof. Dr. Erik C. Böttger (Vorsitz)**
**Prof. Dr. Peter Sander (Leitung der Dissertation)**
**Prof. Dr. Brigitte Berger-Bächi**

**Zürich 2012**

*To my Mum & Dad*

**Acknowledgments**

Once again I learned that a demanding work is the result of many people's efforts. I feel sincerely grateful to those people who made all this possible.

I wish to start thanking my doctoral advisor Prof. Dr. Erik C. Böttger, for giving me the opportunity to carry out my PhD in his Institute, and Prof. Dr. Peter Sander, my supervisor. Peter, I hope I could show you my gratitude for being a wise and patient guide, for teaching me to be an independent scientist, and giving me trust and support. I would like to thank also Prof. Dr. Brigitte Berger-Bächi for joining my PhD committee.

During my project, I was very lucky to collaborate with excellent people, who shared their knowledge with me, and contributed to get done part of my thesis. I refer to Prof. Dr. Massimo Coletta's lab and Prof. Dr. Menico Rizzi's lab, who actively participated in my work, believing in a common project which came out to be very fruitful.

A special thanks to Prof. Dr. Antonio Baici, for the generous help he offered to me and the pleasant time we spent together trying to get into the world of enzyme kinetics.

The proteomic part of my work would have not been possible without the great support of the people of the Functional Genomic Center of the University of Zurich.

Looking at the whole period of the PhD, a group of people has always been on my side: my colleagues. Someone left, someone just arrived, but all in all, they meant much for me. I have to start with the one I am the most grateful to, Beat. Thanks Beat, to have guided me throughout the project, and for teaching me always with enthusiasm. Our students Corrado, Britta, Janine, I can still smile thinking on you, and the fun we had together. Silke and Tom, I warmly thank you for the strong input and motivation you gave me. To the girls group, Giusy, Tanja and Begonia, a huge thank for lighting up my days and giving me a positive vision of life. I am really thankful to the PhD program's team, Celine, Anna, Tina and Koshika, for sharing this experience with me, and spending a great time all together.

To the members of the research group of Peter, Erik and Frank, thanks for contributing to a friendly working environment.

The strength I needed throughout my life, and which played a major role in completing this work, is my family in Rome. To them, I actually don't find words to show my gratitude. My best friends Flami, Lilla and Livia, are part of my family and they represent my mainstay for 15 years. Girls, thanks for your love and unfailing support. Valeria and Massimiliano, who made my experience in Zurich unforgettable, and will always be in my heart, wherever I go. Marina and Roberto, thanks to your warmth I felt at home in Zurich.

It is not easy to find the right place here for the man I love. He would be the first to thank, and in my heart, he is. Reto, every step of this work was done thank to your support, your encouraging words and your optimism.

This work is dedicated to those who gave me happiness and love since I was born, my parents. Mum and Dad, to you my boundless gratitude.

# Content

**Summary**

*Mycobacterium tuberculosis* is the causative agent of one of the world's deadliest diseases, tuberculosis. The obligate pathogen parasites host macrophages by interfering with several cell-host pathways. One of the most peculiar feature of *M. tuberculosis* is the ability to arrest the phagosome maturation process. A new strategy of phagosome maturation arrest dependent on the *M. tuberculosis* zinc metallopeptidase Zmp1 has been recently described. Zmp1 suppresses the activation of a multiprotein complex termed inflammasome, preventing the caspase-1 processing of IL-1β and the subsequent phagosome maturation. Zmp1 is required for virulence of *M. tuberculosis* in mice. The emerging role of Zmp1 as virulence determinant was exploited for generating a vaccine based on the inactivation of *zmp1*. A BCG *zmp1* deletion mutant showed an increased immunogenicity in mice compared to the wild-type. The key role of Zmp1 in mycobacterial pathogenicity and its application in vaccine made of utmost importance the functional characterization of the enzyme. The aims of this study were 1) to provide insights into the structure and biochemical properties of Zmp1 and 2) to generate recombinant live vaccine candidates against tuberculosis based on *zmp1* inactivation.

First, the crystal structure of the *M. tuberculosis* Zmp1 was resolved at 2.6 Å by co-crystallization in complex with the competitive inhibitor phosphoramidon. The Zmp1 structure revealed significant similarities with the human neprilysin and endothelin converting-enzyme 1, identifying Zmp1 as the first prokaryotic member of the M13 endopeptidases. The Zmp1 substrate specificity was investigated by employing a microarray based peptide library. Several bioactive peptides were identified as *in vitro* substrates, including the neuropeptides bradykinin, neurotensin, neuropeptide FF, substance P and apelin. We determined the specific cleavage site within neuropeptides by mass spectrometry, concluding that Zmp1 shares structure, substrate specificity and substrate cleavage pattern with the human neprilysin and endothelin converting enzyme-1. Within the M13 endopeptidase family, Zmp1 represents the first prokaryotic member which is structurally and biochemically characterized.

The second part of this work focuses on the generation of *zmp1*-based vaccine strains with appropriate characteristics for further application in humans, such as sensitivity towards first and second line tuberculosis drugs. BCG *zmp1::aph* and BCG *ureC::hly* are two promising vaccine candidates. The first is in preclinical studies and the second entered the clinical trials. However, both carry antibiotic markers and therefore their use is limited. To overcome these limitations, BCG Δ*zmp1* and BCG *ureC::hly* were generated in the substrain Denmark, without antibiotic resistance markers. Moreover, the combinatorial vaccine strain BCG Denmark Δ*zmp1 ureC::hly* was generated as the combination of the properties of the single mutants might further improve the protective efficacy and safety profile of the vaccine.

**Zusammenfassung**

*Mycobacterium tuberculosis* ist der Erreger der Tuberkulose, eine der tödlichsten Infektionskrankheiten weltweit. Der Krankheitserreger parasitiert in Wirtsmakrophagen indem er zelluläre Abwehrmechanismen inhibiert. Eine der bemerkenswertesten Eigenschaften von *M. tuberculosis* ist dessen Fähigkeit die Phagosomenreifung zu inhibieren. Ein neuartiger Mechanismus der Inhibition der Phagosomenreifung, basierend auf der *M. tuberculosis* Zink-Metallopeptidase Zmp1, wurde erst kürzlich beschrieben. Zmp1 unterdrückt die Aktivierung eines Multiproteinkomplexes, Inflammasom genannt, und damit die Caspase-1-abhängige pro-IL-1β Prozessierung. Matures IL-1β stimuliert die Reifung von Phagosomen. Eine *M. tuberculosis zmp1* Deletionsmutante ist attenuiert. Aufgrund der zentralen Bedeutung von Zmp1 wurde eine *Mycobacterium bovis* BCG *zmp1* Mutante als potentieller Tuberkuloseimpfstoff generiert. Diese BCG *zmp1* Deletionsmutante wies im Vergleich zu BCG eine erhöhte Immunogenität in Mäusen auf. Die Schlüsselrolle von Zmp1 in der Pathogenität von Mykobakterien sowie die Eigenschaften eines *zmp1*-basierten Impfstammes lassen eine funktionelle Charakterisierung dieses Enzyms äusserst interessant erscheinen. Ziele dieser Studien waren 1) die strukturelle und biochemische Charakterisierung von Zmp1 und 2) die Generierung verbesserter rekombinanter *zmp1*-basierter Lebendimpfstoffe gegen die Tuberkulose.
Zunächst wurde die Kristallstruktur von *M. tuberculosis* Zmp1 als Komplex mit dem kompetitiven Inhibitor Phosphoramidon mit einer Auflösung von 2.6 Å ermittelt. Die Struktur von Zmp1 zeigte dabei signifikante Ähnlichkeit mit humanem Neprilysin und *Endothelin-converting enzyme-1* (ECE-1). Zmp1 ist der erste Vertreter der M13 Endopeptidasen, dessen Struktur bestimmt wurde. Die Substratspezifität von Zmp1 wurde mit Hilfe einer Micro-Peptidbank bestimmt. Mehrere bioaktive Peptide konnten dabei als *in vitro* Substrate identifiziert werden, darunter die Neuropeptide Bradykinin, Neurotensin, Neuropeptid FF, *Substance P* und Apelin. Die Schnittstellen innerhalb der Neuropeptide wurden mittels Massenspektrometrie ermittelt und ähnelten denjenigen von Neprilysin und ECE-1. Somit ähnelt Zmp1 sowohl bezüglich der Struktur als auch bezüglich der Substratspezifität und Substratschnittstellen den humanen Peptidasen Neprilysin und ECE-1. Innerhalb der

M13 Endopeptidase Familie repräsentiert *M. tuberculosis* Zmp1 damit den ersten prokaryontischen Vertreter, welcher strukturell und biochemisch charakterisiert wurde.

Der zweite Teil dieser Arbeit fokussiert auf der Weiterentwicklung von *zmp1*-basierten Impfstämmen. Für die Anwendung am Menschen ist insbesondere deren Sensitivität gegenüber *First-* und *Secondline* Tuberkulostatika wichtig. BCG *zmp1::aph* und BCG *ureC::hly* sind zwei vielversprechende Impfstoffkandidaten. Ersterer befindet sich in vor-klinischen Studien, während sich der Zweite bereits in klinischen Studien befindet. Beide Stämme tragen jedoch Antibiotikaresistenzmarker, was ihre Verwendung einschränkt. Um diese Limitierung zu umgehen wurden in dem BCG Substamm Denmark die Allele Δ*zmp1* und *ureC::hly* ohne jedweden Antibiotikaresistenzmarker eingebracht. Darüber hinaus wurde ein kombinatorischer Stamm, BCG Denmark Δ*zmp1 ureC::hly* generiert. Durch die Kombination der Eigenschaften der Parentalstämme dürfte eine weitere Verbesserung der Protektivität und Sicherheit erreicht werden.

## Abbreviations

| | |
|---|---|
| Ag85 | Antigen 85 |
| Amp | ampicillin |
| BCG | Bacille Calmette-Guérin |
| bp | basepair |
| CFU | colony forming units |
| Δ (Delta) | deletion / gene knock out |
| ESAT-6 | 6 kDa early secretory antigenic target *esxA* |
| Hyg | hygromycin |
| Kan | kanamycin |
| kDa | kilodalton |
| KO | knock out |
| MDR-TB | multidrug-resistant *M. tuberculosis* |
| Mtb | *M. tuberculosis* |
| MW | molecular weight |
| OD | optical density |
| ORF | open reading frame |
| sco | single cross-over |
| TB | tuberculosis |
| wt | wildtype |
| XDR-TB | extensively drug-resistant *M. tuberculosis* |

# INTRODUCTION

## 1. Tuberculosis. History, epidemiology and therapy

*In the first papers concerning the aetiology of tuberculosis I have already indicated the dangers arising from the spread of the bacilli-containing excretions of consumptives, and have urged moreover that prophylactic measures should be taken against the contagious disease. But my words have been unheeded. It was still too early, and because of this they still could not meet with full understanding. It shared the fate of so many similar cases in medicine, where a long time has also been necessary before old prejudices were overcome and the new facts were acknowledged to be correct by the physicians.*

Robert Koch
"The current state of the struggle against tuberculosis", Nobel Lecture (12 Dec 1905)

One century after Robert Koch's Nobel lecture, tuberculosis (TB) still remains a major global health threat. Tuberculosis has plagued mankind worldwide for thousands of years. More than 2000 years ago in ancient Greece, Hippocrates called the most common cause of illness in his time "phthisis" (Hippocrates, *Aphorisms*). Over the centuries since Hippocrates, tuberculosis has been known as a major scourge of the human species. The disease was named the Great White Plague in the 17$^{th}$ century in Europe, where tuberculosis was epidemic and the principal cause of death for two-hundred years (Dubos, 1952, *The White Plague: Tuberculosis, Man, and Society).* Malnutrition, high population density as well as poor sanitary conditions in the European and North American cities created a fruitful field for spreading of the disease. The historic menace of the "White Plague" continued for many centuries because of poor understanding of the disease and unappropriate medical tools to fight it. Tuberculosis still remains a disease of poverty strongly interconnected with undernutrition and overcrowding (Lawn and Zumla, 2011). Most cases occur in Asia

and Africa, with smaller proportions of cases in European region, in the eastern Mediterrean and the Americas. The devastating effect of HIV on susceptibility to tuberculosis has dramatically increased the tuberculosis incidence rate in many high HIV prevalence countries in sub-Saharan Africa (about four of every five cases) (WHO, 2010). In fact, the risk of developing tuberculosis for HIV infected people was evaluated to be more than 20-times greater than that of non-infected individuals (WHO, 2009).

The incidence of TB in 2010 is worrisome; 8.8 million new cases have been estimated, and 1.4 million people died from TB. Overall, one-third of the world's population is currently infected with the aetiologic agent *Mycobacterium tuberculosis* (Mtb), but only 5-10% of infected people (who are not infected with HIV) will develop active disease in their lifetime (WHO, 2010). Indeed, most immuno-competent individuals infected with *M .tuberculosis* either eliminate the pathogen or co-exist with it in a latent state. Latent tuberculosis is a clinical disorder in infected individuals in whom the host immune system keeps under control the bacterial multiplication such that the infection remains asymptomatic. In those individuals, the infection can reactivate years later or even decades, resulting in active and contagious tuberculosis. The risk of reactivation is increased in immuno-compromised persons, such as HIV infected individuals (Algood *et al.*, 2003).

*M. tuberculosis* is an obligate pathogen, able to infect several animal species, but human beings are its principal host. As aerobic bacillus, it grows preferentially in tissue with high oxygen content, such as lungs. The complex and not fully understood strategy of survival of *M. tuberculosis* in the host represents its key of success. *M. tuberculosis* hides into macrophages, targeting therefore the primary host defence against microbial invasion. The successful parasitization of macrophages results in the multiplication of the bacteria while the immune system is silenced. The intracellular survival is guaranteed by multiple factors. The thick mycobacterial cell wall represents a physical barrier to mechanical and chemical stress, such as the hydrolytic enzymes within the macrophages. *M. tuberculosis* virulence factors comprise a real arsenal for the defence against the host attacks, such as the mycobacterial enzymes that detoxify reactive radicals. The most striking feature of

pathogenic mycobacteria is the ability to block the normal process of phagosome maturation, resulting in the growth in a non-acidified intracellular compartment.

The WHO-recommended chemotherapeutic regimen against tuberculosis consists of four out of five first line drugs (rifampicin, isoniazid, pyrazinamide, ethambutol and streptomycin) for initial two months, and a following treatment with rifampicin and isoniazid for the next four months. This long-lasting therapy is designed to kill the replicating bacteria in the first phase with usage of multiple drugs and the next four months treatment is supposed to eradicate the remaining non-replicating bacteria. Nevertheless, many factors thwart the success of the therapy. The granuloma, because of its highly differentiated internal structure, offers multiple micro-enviroments for Mtb. Hence, it is populated by bacterial sub-populations in different growth phases and therefore with diverse metabolic activities. The drugs currently used preferentially target replicating organisms, so that the nonreplicative sub-populations are protected by innate resistance mechanisms. The innate resistance of nonreplicating Mtb contributes substantially to the protracted chemotherapeutic treatment (Russell *et al.*, 2010).

The duration, high cost and drug-related adverse effects of the therapy are often the reason for patient noncompliance, which, worsened by inadequate health-care oversight, results in increasing proliferation of drug-resistant strains. MDR-TB (multidrug-resistant tuberculosis) is a specific form of TB caused by resistant Mtb to at least isoniazid and rifampicin, the two most powerful anti-TB drugs. MDR-TB strains further resistant to fluoroquinolone and an injectable second-line drug (kanamycin, amikacin and/or capreomycin) are called XDR-TB (extensively drug-resistant tuberculosis). While drug-susceptible TB can be cured within six months, MDR-TB requires extensive chemotherapy (with drugs which are more expensive and have severe side effects) for up to two years. Instead, for XDR-TB cases treatment options are seriously limited. It was estimated that 440 000 people developed MDR-TB worldwide in 2008 and that a third of them died. Almost 50% of MDR-TB cases worldwide are estimated to occur in China and India (WHO's Multidrug and Extensively Drug-Resistant Tuberculosis: 2010 Global Report on Surveillance and Response). The growing global burden of multidrug-resistant tuberculosis urges the need for development of fast-acting medications and new effective vaccines.

## 2. Microbiology of *M. tuberculosis*

*Mycobacterium* is a genus of Actinobacteria which includes more than 130 species. The genus encompasses a number of medically important species, including the aetiological agent of tuberculosis *M. tuberculosis* and the causative agent of leprosy *M. leprae*. Species relevant for their applications are *M. bovis* BCG, the vaccine strain used against tuberculosis, and *M. smegmatis*, a non-pathogenic mycobacterial model organism. Mycobacteria are straight or slightly curved nonmotile rods (approximately 0.2–0.6 µm wide by 1.0–10 µm long), nonsporeformers. Two remarkable features define mycobacteria. First, they have a thick cell wall enriched in uncommonly long-chain fatty acids with 60 to 90 carbons, known as "mycolic acids", attached to the cell wall through the polysaccharide arabinogalactan. The mycolic acids layer give the colonies their characteristic waxy appearance and the cells a tendency to clump and resist dispersion. The waxy nature of the cell envelope makes the bacteria weakly stainable with Gram's method dyes, though the bacteria are categorized as Gram positive. Mycobacterial mycolic acids form, however, strong complexes with some chemicals, as carbol fuchsin dye, that leave them resistant to decolorization with acid alcohol. For that reason mycobacteria are defined acid-alcohol fast. In addition, the lipid coat confers to mycobacteria a strong resistance to many antiseptic solutions and antibiotics. Second, mycobacteria grow relatively slowly with generation times ranging from 2 hours for *M. smegmatis* (classified therefore as fast growing mycobacteria) to 12 days for *M. leprae* (slow growing mycobacteria). *M. tuberculosis* belongs to the group of slow growing mycobacteria, with a generation time of 20 to 24 hours. The genus *Mycobacterium* shares the production of mycolic acids and an unusually high genomic DNA GC content (62–70%) with two closely related genera, *Nocardia* and *Corynebacterium*, within the order *Actinomycetales*. The complete genome sequence of the *M. tuberculosis* laboratory strain H37Rv was elucidated in 1998 (Cole *et al.*, 1998). The genome sequence comprises 4.4 Mbp, and contains 3,924 open reading frames. The genome is rich in repetitive DNA sequences, particularly insertion sequences in intergenic or non-coding regions, often near tRNA genes. The mycobacterial genome is composed by genes needed for adaptation of the bacterium to a range of environmental conditions,

from growth in a rich broth to survival in host macrophages. Moreover, some components of several anaerobic phosphorylative electron transport chains were found, conferring to the bacterium the ability to survive also in anaerobic environment, such as the core of granuloma. The genome also possesses an extraordinary large number of genes involved in lipid metabolism: ~250 distinct enzymes involved in complex fatty acid biosynthesis and degradation.

The architecture of the mycobacterial cell envelope is peculiar and rather complex. It includes the plasma membrane, the periplasmic space, the mycomembrane and the capsular layer. Such unique structure is placed in between the Gram positive and Gram negative cell envelope structures. The cytoplasmic membrane is encapsulated by a layer of peptidoglycan, which peculiarly contains N-glycolylmuramic acid instead of N-acetylmuramic acid (Brennan and Nikaido, 1995). The peptidoglycan is connected through a phosphodiester link to a second component of the periplasmic layer, the arabinogalactan. The mycolic acids are bound to arabinogalactan by ester bonds, forming the mycomembrane. By hydrophobic interactions, the mycolic acids interact with free glycopeptidolipids and free glycolipids, creating an outer membrane-like structure, which is followed by a capsular layer. A major component associated to the mycobacterial cell wall is the immunogenic lipoarabinomannan (LAM), which is attached to the cytoplasmic membrane by a phosphatidylinositol anchor. *M. tuberculosis* LAM are mannose capped (ManLAM), whereas nonpathogenic species lack the glyco-modification (Vercellone *et al.*, 1998). The mannose-capped lipoarabinomannan have been shown to be the key effector in the process of phagosome maturation arrest (Fratti *et al.*, 2001; Fratti *et al.*, 2003).

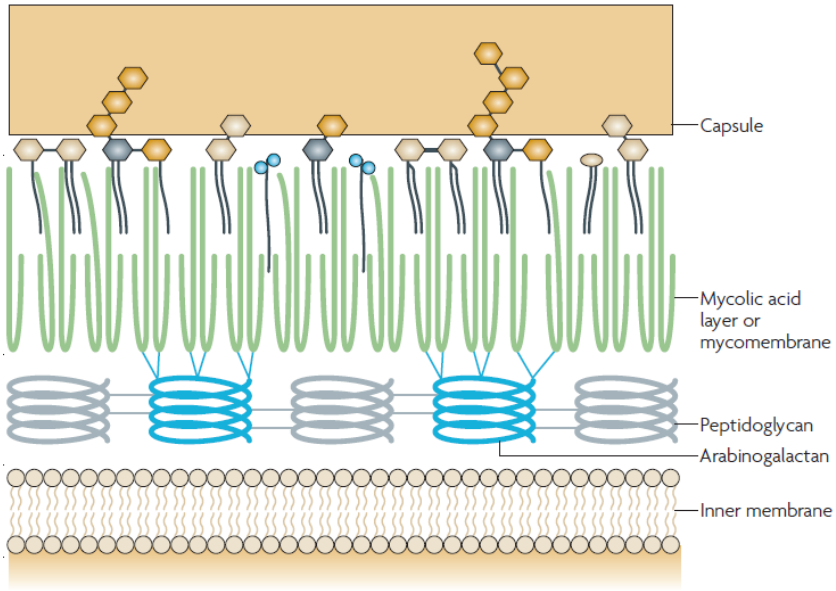



Fig.1 Schematic representation of the architecture of the mycobacterial cell wall (Abdallah *et al.*, 2007).

## 3. Pathogenesis of tuberculosis

Infection with *M. tuberculosis* follows a well-established pattern of events (Fig.2). The infectious bacilli are inhaled as droplet nuclei from the atmosphere. These droplets remain airborne for several hours, and the minimum infection dose ranges from a single bacterium upward (Russell *et al.*, 2010). In the lung the bacteria are phagocytosed by alveolar macrophages, inducing a localized inflammatory response that leads to recruitment of mononuclear cells from neighboring blood vessels, providing fresh host cells to infect. These cells are the starting point for building the granuloma, which is the pathologic feature of tuberculosis. In the early stage, granuloma is an amorphous mass composed of macrophages, neutrophils and monocytes; macrophages, however, undergo to differentiation into multinucleated giant cells, epithelioid macrophages and foamy macrophages. When lymphocytes arrive and the acquired immune response is initiated, the granuloma develops into an organized, and stratified structure, surrounded by a fibrous cuff which clearly confines the mass. The active bacterial replication lasts for two to three weeks; afterwards, the appearance of Mtb-specific lymphocytes symbolizes the end of this phase and the beginning of the "containment" state. During that phase, the granuloma is extensively vascularized, the cells are continuously recruited and the

bacterial count is relatively stable. The disease progression is then associated with decreasing vascularization and thickening of the fibrous sheath, along with the increase of number of foamy macrophages mainly responsible for the typical caseous areas in the center of the granuloma. The late stage of the infection is characterized by hypoxic condition inside the granuloma, which induces a state of non replication of bacteria. Active granulomas exhibit extensive pathology, and their rupture is responsible for spilling of thousand of viable, infectious bacilli into the airways. At this step, spreading of the infection occurs by coughing. Containment of rupture usually fails when the host immune system is compromised, and this may correspond to malnutrition, aging or co-infection with HIV. In general, any condition which affects the number or function of CD4 T cells represents a high risk for activation of the disease (Russell, 2007; Russell *et al.*, 2010).

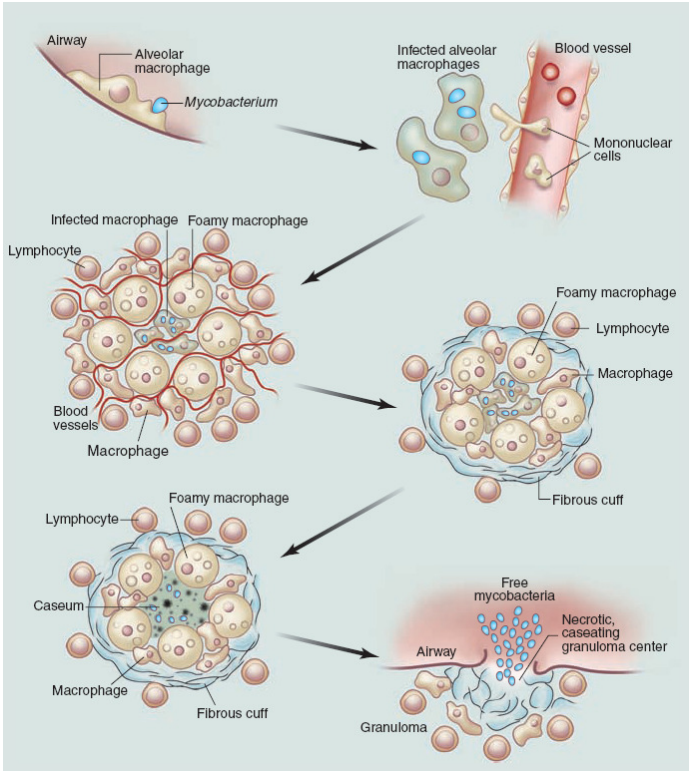



Fig.2 The progression of *M. tuberculosis* pathology (Russell *et al.*, 2010).

Granuloma play a contradictory role in tuberculosis pathology. Granuloma formation functions as a host defense mechanism, which contains the infection site. Indeed, macrophages are able to kill mycobacteria, and control their growth. Although *M. tuberculosis* inhibits apoptosis and immune responses, it is still capable of basal CD4 T cells stimulation, and in minor extent CD8 T cells stimulation

(Kaufmann, 2004). In fact, in HIV infected individuals, granulomas are absent, resulting in an uncontrolled Mtb dissemination (Lawn *et al.*, 2002). On the other hand, granuloma confer many advantages to *M. tuberculosis*. Pathogenic mycobacteria promote cellular recruitment to the granuloma, and newly recruited macrophages provide their primary growth niche (Davis and Ramakrishnan, 2009). Inside the granuloma, mycobacteria might evade massive and effective innate immune mechanisms thereby delaying the start of an adaptative immune response. Moreover, such dense structure shields themself from therapeutic concentration of anti-tuberculosis drugs, potentially promoting the emergence of drug-resistant strains (Lawn and Zumla, 2011). From this point of view, granuloma formation might be seen as a part of the mycobacteria directed virulence program.

Tuberculosis affects predominantly the lungs, and the most common clinical manifestation of tuberculosis is pulmonary disease. Nevertheless, about 15% of the patients with active tuberculosis, develop miliary or extra-pulmonary TB. The bacteria, in fact, can disseminate to diverse parts of the body through invasion of blood stream. The interested organs are lymph nodes, liver, pleura, spleen, bones and joints. The most serious clinical manifestation of tuberculosis is the involvement of the central nervous system, resulting in tuberculous meningitis, which is fatal in almost all cases without a prompt chemotherapy (Frieden *et al.*, 2003). Disseminated tuberculosis is characterized by involvement of many organs simultaneously and it results from primary progressive disease or reactivation of latent infection. General symptoms caused by TB are persistent cough, fever, weight loss, night sweat, shortness of breath and chest pain. However, in the early stage of disease, mostly because of macrophages containment work, symptoms might be absent (Ducati *et al.*, 2006).

## 4. The life of *M. tuberculosis* inside the phagosome

The successful parasitization of the macrophage by *M. tuberculosis* involves the manipulation of the host signalling pathways, causing the impairment of its natural main function: innate and adaptative immune response effector. Pathogenic mycobacteria, unlike non-pathogenic species, are capable of inhibition of several host

processes in order to guarantee the survival and multiplication inside the host. Four main processes are affected by Mtb: phagosome maturation, antigen presentation, apoptosis and stimulation of bactericidal responses (Fig.3A). Pathogenic mycobacteria, in contrast to non-pathogenic mycobacteria, disrupt the mitogen-activated protein kinase (MAPK) pathway inhibiting the activation of p38 and ERK1/2. As consequence, the production of pro-inflammatory cytokines and chemokines, which are needed to induce a cellular innate immune response, is impaired (Koul *et al.*, 2004). The corruption of the host apoptotic pathway is a strategy delivered by Mtb to avoid cross-priming, which leads to CD8 T cell stimulation. The mycobacterial virulence factor Man-LAM promotes the phosphorylation of the apoptotic protein Bad through the protein kinase B Akt, leading to its dissociation from the anti-apoptotic protein Bcl-2, and its consequent anti-apoptotic effects in infected cells (Maiti *et al.*, 2001). Mtb also limits macrophage apoptosis by inducing the production of the immuno-suppressive cytokine interleukin-10 (IL-10), which in turn blocks the synthesis of the potent stimulator of apoptosis, tumour-necrosis factor-α (TNF-α), in infected macrophages (Balcewicz-Sablinska *et al.*, 1998).

The most peculiar feature of *M. tuberculosis* is the ability to arrest the phagosome maturation process. Phagosome maturation involves a sequence of fusion events with various vesicles, by which the nascent phagosome maturates into a specialized acidic compartment with microbicidal properties, the phagolysosome. After bacterial ingestion, the phagosome acquires markers, such as the small GTPase Rab5, that recruits phosphatidylinositol-3-kinase (PI3K). This kinase produces a molecule, phosphatidylinositol-3-phosphate (PI3P), necessary for recruitment of early endosomal antigen (EEA1). EEA1 is the key effector to allow the maturation process through the fusion of phagosomes with late endosomes. It serves as an organelle tethering molecule by bridging membranes destined for fusion (Vergne *et al.*, 2004a). During maturation, the phagosome loses Rab5 and acquires another GTPase, Rab7, and following fusions with lysosomal vesicles equip the compartment with the lysosome-associated membrane protein 1 (LAMP1), acid hydrolysases, such as cathepsin D, and the vacuolar proton-ATPase which acidifies the phagolysosome. The whole process is $Ca^{2+}$ sensitive. In fact, in the early steps, an increased

intracellular concentration of $Ca^{2+}$ activates calmodulin and calmodulin-dependent protein kinase CaMKII, which are responsible for recruiting PI3K. *M. tuberculosis* evolved a mechanism to arrest phagosome maturation playing on its critical point, $Ca^{2+}$ concentration. The pathogen inhibits cytosolic $Ca^{2+}$ rise through the lipid effector Man-LAM, which is able to block ionophore-induced increases in $Ca^{2+}$concentration in macrophages (Rojas *et al.*, 2000). As Man-LAM inhibits the increase in cytosolic $Ca^{2+}$ concentration, it impairs the $Ca^{2+}$ /calmodulin recruitment of PI3K to the phagosomes and thereby further obstructs phagosomal maturation. Mycobacterial phagosomes therefore do not recruit EEA1, which, as mentioned above, is essential for the fusion of the phagosome with lysosomal vesicles. The incomplete maturation of the *M. tuberculosis* vacuole confers unique characteristics to this environment: absence of mature hydrolases, limited acidification to pH 6.4 (phagolysosome pH is usually 4.8) and active trafficking with early endosomes for acquisition of nutrients (Russell, 2001; Vergne *et al.*, 2004a) (Fig.3B). It is clear that pathogenic mycobacteria arrest phagosome maturation. However, the molecular details have yet to be elucidated. So far, several mechanisms have been suggested, taking to the conclusion that multiple mycobacterial effectors are involved in the corruption of phagosome trafficking (Pethe *et al.*, 2004). Surface lipids such as Man-LAM and cord factor (a mycobacterial trehalose dimycolate) have been shown to inhibit $Ca^{2+}$ -induced vesicular fusion (Fratti *et al.*, 2003; Indrigo *et al.*, 2003; Vergne *et al.*, 2003), and the phosphatidylinositol mannoside (PIM) stimulates fusion between phagosome and early endosomes, that are a source of nutrients (Vergne *et al.*, 2004b). Moreover, some mycobacterial proteins revealed a key role in this process, as the bacterial serine/threonine kinase PknG (Walburger *et al.*, 2004), the mycobacterial urease UreC, which produces ammonia that blocks vacuole acidification (Reyrat *et al.*, 1996), the acid phosphatase SapM (Vergne *et al.*, 2005) and the zinc metallopeptidase Zmp1 (Master *et al.*, 2008). It is well established that macrophage activation with IFN-γ prior to infection with *M. tuberculosis* enables the host cell to overcome the blockage in phagosome maturation and deliver the bacterium to an acidic, hydrolytically competent phagolysosome (Schaible *et al.*, 1998). The mechanism behind was suggested to rely on a small GTP-binding protein, LRG-47, which is selectively upregulated upon IFN-γ activation, and it facilitates the delivery

of vacuolar $H^+$ ATPase subunits to the phagosome, supporting the phagosome acidification (MacMicking *et al.*, 2003).

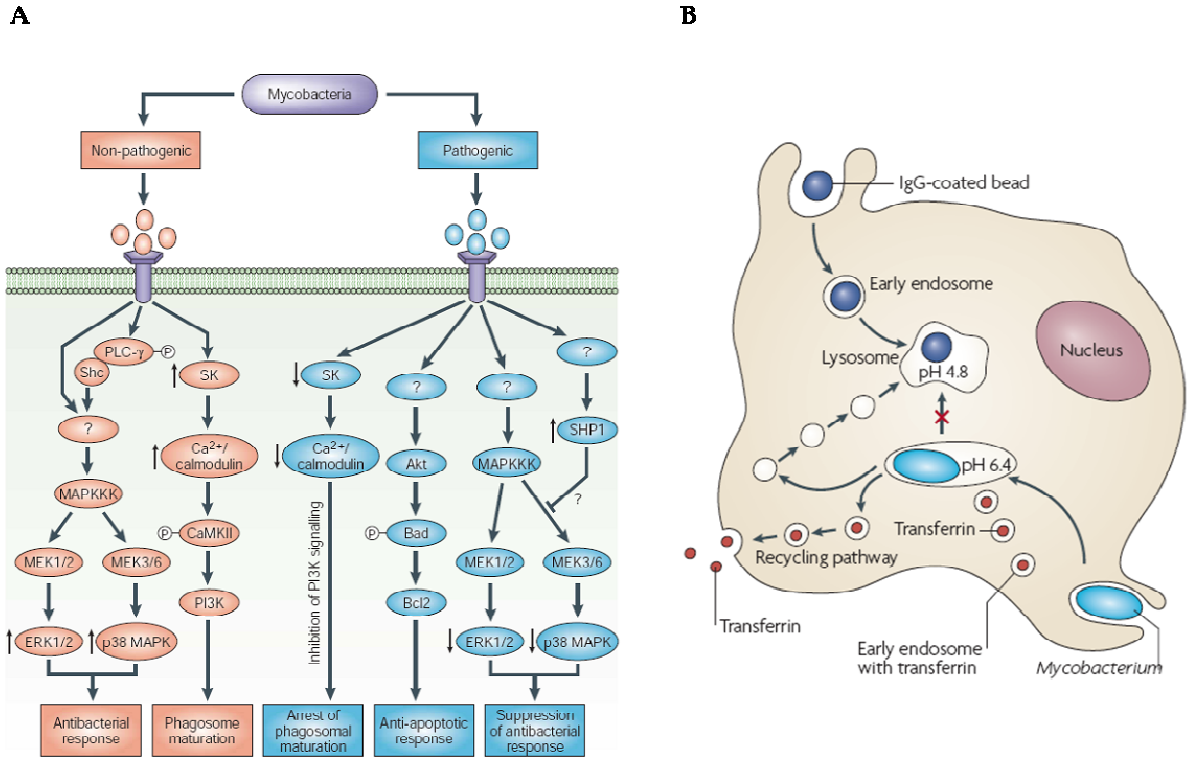



Fig.3 (A) Representation of the host signalling pathways regulated by pathogenic and non-pathogenic mycobacteria (Koul *et al.*, 2004). (B) The vacuole containing *M. tuberculosis* appears as a not hostile environment with respect to pH, hydrolytic activity and nutrient supply. Phagosomes containing IgG-coated beads acidify rapidly to pH lower than 5 because of fusion with lysosomes (Russell, 2007).

## 5. Zinc metallopeptidases. An overview from mammals to bacteria

Zinc metallopeptidases have physiological roles in the life cycle of the organisms, and they are essential factors for homeostatic control in both eukaryotes and prokaryotes. Several opportunistic pathogens use zinc metallopeptidases as toxic factors to the host. These toxic peptidases show a variety of pathological actions, because of their versatile proteolytic activity towards many kinds of host proteins, such as plasma proteins and structural component of the tissues. With the exception of a few cases, the precise role of zinc metalloproteases in bacterial pathogenesis has

not yet been elucidated. The striking virulence of some pathogens relies solely on zinc metallopeptidase. Glaring examples are the metallopeptidase from *Vibrio vulnificus* which, enhancing the vascular permeability by stimulating the production of inflammatory mediators histamine and bradykinin, causes serious septicemia and edematous skin lesions in humans (Miyoshi *et al.*, 1994); *Vibrio cholera* secretes a metallopeptidase which activates cholera toxin, a prototype A-B subunit toxin, through digestion of the A subunit (Booth *et al.*, 1984), and a metallopeptidase from the *Vibrio cholera* serovar O1, the causative agent of pandemic cholera, accelerates the bacterial attachment to intestinal epithelial cells through digestion of the small intestine mucosa, thereby destroying the protective barrier and providing nutrients to the bacterial population (Ichinose *et al.*, 1994). Beside the impact of zinc metallopeptidases in enteropathogenicity and vascular permeability, necrotic tissue damage represents the most classic pathogenic mechanism of bacterial peptidase. Pseudolysin, an extracellullar elastase of *Pseudomonas aeruginosa*, destroys tissue and damages cell functions, has hemorrhagic activity, muscle damaging effects, and causes chronic ulcers by degradation of human wound fluids and human skin proteins (Adekoya and Sylte, 2009). Among the bacterial metallopeptidases with the most devastating effect belongs the lethal factor of *Bacillus anthracis*. The peptidase is a part of the three component anthrax toxin, and it works inactivating key signaling molecules, such as mitogen-activated protein kinase kinases (MAPKK), to ultimately cause cell death (Panchal *et al.*, 2004).

Zinc metallopeptidases are classified into four groups, namely DD-carboxypeptidase, carboxypeptidases, zincins and inverzincin (Fig.4). The 'zincins superfamily' is characterized by the zinc-binding motif HEXXH. This superfamily is subdivided in seven families, and the most representative bacterial ones are three: thermolysin having *Bacillus thermoproteolytics* metallopeptidase as prototype enzyme, serralysin with the metallopeptidase from *Serratia marcescens* as prototype, and neurotoxin family represented by *Clostridium botulinum* and *Clostridium tetani* metalloenzymes (Miyoshi and Shinoda, 2000). In the HEXXH motif, the two histidines are zinc ligands and the glutamic acid has catalytic function as it polarizes the water molecule that attacks the peptide bond of the substrate (Fig.5). The zinc atom is tetrahedrally-coordinated by a third ligand which is usually a residue of glutamic acid, and a

fourth ligand that is a molecule of water that becomes activated and mediates the nucleophilic attack on the scissile peptide bond (MEROPS).

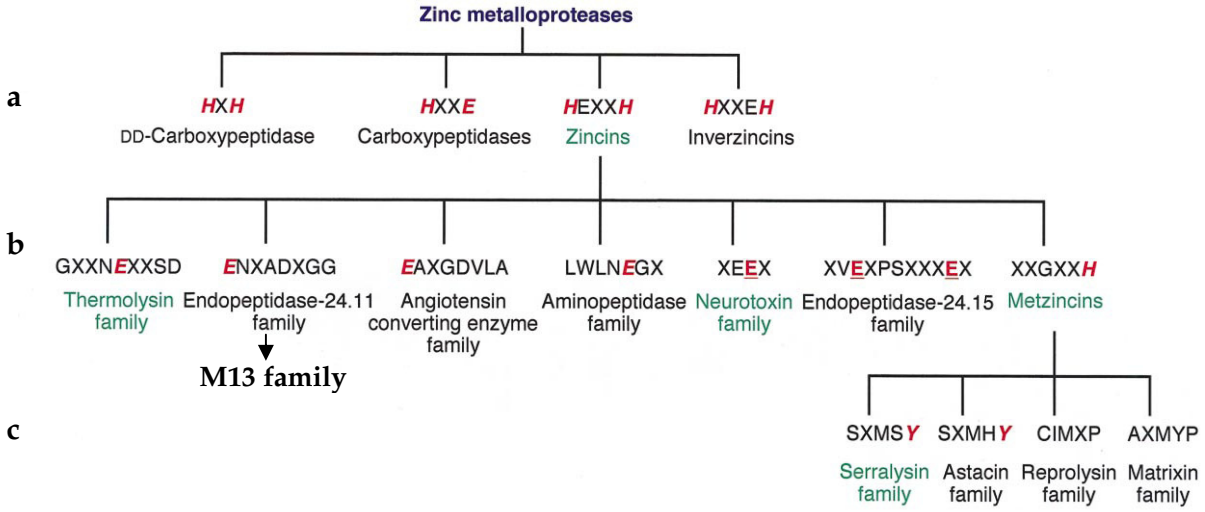



Fig.4 Zinc metallopeptidases families classified on the base of the sequence around the zinc-binding residues. Letters in italic represent the identified zinc ligands; underlined red letters are putative zinc ligands; X stands for any amino acid. Residues in the line **a** correspond to the first and second zinc ligands; residues in the line **b** to the third ligand and those in the line **c** to the putative fifth ligand (Miyoshi and Shinoda, 2000).

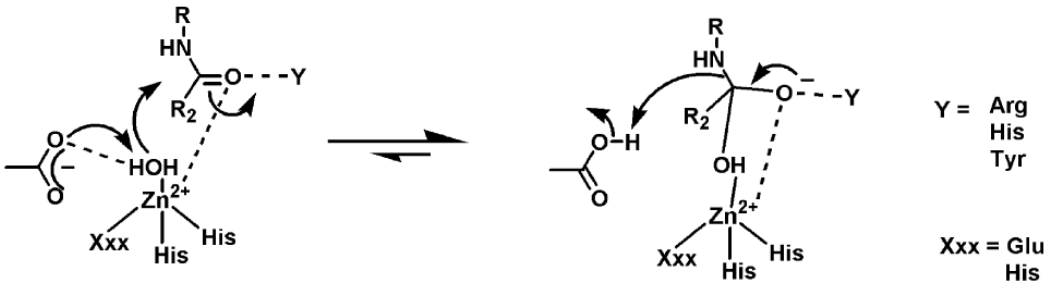



Fig.5 Scheme representing the catalytic mechanism of zinc metallopeptidases (Hernick and Fierke, 2005).

The endopeptidase-24.11 family belongs to the 'zincins superfamily', and it deserves a closer look because of the physiological importance of its members. The family was named in MEROPS database as M13 metallopeptidase family, and the prototype is neprilysin (NEP). The neprilysin family is a large group of medically and developmentally important enzymes. Indeed, the members of this family are

involved in the metabolism of a number of regulatory peptides of the mammalian nervous, cardiovascular, immune and inflammatory system. They target small peptides (not longer than 40 residues) such as tachykinins, enkephalins, natriuretic and chemotactic peptides. The majority of M13 peptidases described so far exhibits a strong preference for cleaving the amino terminal bond of hydrophobic residues (Bland *et al.*, 2008). The prototype neprilysin is the best characterized member, biochemically and structurally investigated. Neprilysin is widely distributed in mammalian tissues and has important role in inactivating signaling peptides. At nervous system level, it degrades endogenously released enkephalins, thereby regulating opioid peptide action, and other neuropeptides of the tachykinin family such as substance P (Turner *et al.*, 2001). Neprilysin modulates blood pressure by physiological degradation of hormone peptides such as cardiac hormone natriuretic peptide (ANP), bradykinin and endothelin (Oefner *et al.*, 2000). NEP takes part also in various cancers, due to its ability to turn off mitogenic peptide signals, mediated by bombesin-like peptides or endothelins (Turner *et al.*, 2001). NEP is potentially an important drug target, and there is a great interest in developing novel inhibitors of NEP for application in clinical treatment as analgesics or anti-hypertensive agents. Phosphoramidon is one of the most selective and potent inhibitor of M13 peptidases in general and NEP in particular. Recent reports showing that brain NEP can degrade the neurotoxic β-amyloid peptide involved in Alzheimer's disease, raises new concerns about the chronic blockade of NEP activity, at least in CNS, by using a potent inhibitor as phosphoramidon. In fact, the NEP inhibition results in pathological accumulation of amyloid *in vivo* (Iwata *et al.*, 2000). Thus, it appears that a critical balance of NEP is needed to guarantee homeostasis of peptide signaling events in order to avoid pathological states.

Other members of the M13 family are endothelin-converting enzymes (ECE-1 and ECE-2), the erythrocyte surface antigen KELL and the phosphate-regulating neutral endopeptidase PHEX. ECE-1 and NEP were discovered on the basis of their activities, whereas all the other members of the family were identified by their sequence similarity. The overall homology between NEP and ECE-1, ECE-2, KELL and PHEX is 35 %, 30 %, 31 %, and 25 %, respectively, but increases to 51 %, 44 %, 43 %, and 39% if only the 250 C-terminal residues are considered (Oefner *et al.*, 2000).

The high degree of similarity suggests that these peptidases have common origin and a similar fold. Their enzymatic activities were investigated over the past years, but so far only the crystal structures of NEP and ECE-1 both complexed with phosphoramidon have been resolved (Oefner *et al.*, 2000; Schulz *et al.*, 2009). The M13 peptidases are type II integral membrane proteins, and they all consist of a short N-terminal cytoplasmatic domain, followed by a single transmembrane helix, and a large C-terminal extracellular domain that contains the active site. Another common structural feature includes the ten conserved cysteine residues in the extracellular domain, four of which are clustered close to the transmembrane anchor. They form intrachain disulfide bridges which confer stability to both structure and function (Tam *et al.*, 1985). The structure of the extracellular domain of NEP is characterized by two domains that enclose a large central cavity containing the active site. The larger N-terminal domain, which shows high similarity with thermolysin, includes the critical residues for zinc coordination and catalysis. The C-terminal domain restricts the active site access and thus functions as molecular sieve (Oefner *et al.*, 2000). The two-domain structure of NEP, with a peculiar substrate size-restricting region, reveals its nature as oligopeptidase (the majority of the natural substrates of NEP are restricted to about 3000 Da), whereas thermolysin, missing that domain, cleaves protein substrates (Fig.6).

The bacterial homologues of the M13 family have not yet been characterized. In mycobacteria little is known about metallopeptidases and their potential role in virulence. TubercuList and MEROPS databases report only six annotated metallopeptidases in *M. tuberculosis*, on the basis of the common HEXXH motif. These are the zinc metallopeptidase Zmp1 (Rv0198c), the aminopeptidase N (Rv2467), the membrane-bound protease FtsH (Rv3610c), the protease heat shock protein X (Rv0563), the hypothetical protein encoded by the ORF Rv1977 and the membrane-bound zinc metallopeptidase Rip (Rv2869c). Both zinc metallopeptidases Zmp1 and Rip have been reported as important virulence determinants. Rip, a non-essential intramembrane cleaving protease, regulates cell envelope lipid composition in *M. tuberculosis* by cleaving three transmembrane anti-sigma factors which control multiple downstream pathways involved in lipid biosynthesis (Sklar *et al.*, 2010). Rip

is required for both *M. tuberculosis* growth and persistence in mice (Makinoshima and Glickman, 2005).

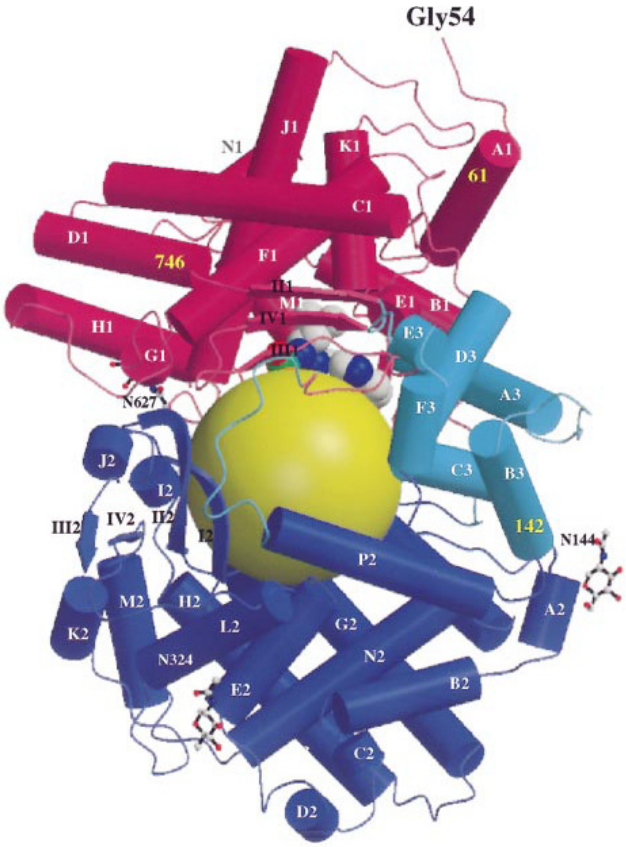



Fig.6 Ribbon plot of soluble NEP depicting the volume of the active site cavity. The large N-terminal catalytic domain is colored in violet and the smaller C-terminal domain in blue. The three inter-domain linker fragments are represented in cyan. The inhibitor phosphoramidon is nestled in the interior of the cavity of the large catalytic domain (Oefner *et al.*, 2000).

Recently, a new strategy of mycobacterial phagosome maturation arrest relying on the *M. tuberculosis* zinc metallopeptidase Zmp1 has been described (Master *et al.*, 2008). The evidences of the involvement of Zmp1 in virulence and survival in macrophages were provided by using mycobacteria deficient in Zmp1. *In vivo* experiments showed that Zmp1 is necessary for full virulence in mice, since Mtb knockout mutant displayed reduced survival in mice. Moreover, phagosomes harboring Zmp1 mutant mycobacteria undergo enhanced maturation into phagolysosomes, suggesting that Zmp1 is required to maintain mycobacterial phagosome maturation arrest. The ability of mycobacteria to block phagosome

maturation relies on the Zmp1-dependent suppression of inflammasome activation. By the Zmp1-dependent inhibition of caspase-1 activation, mycobacteria prevent IL-1β production and subsequently phagosome maturation. The exact mechanism of interaction between Zmp1 and the inflammasome is still unknown. Undoubtedly, Zmp1 plays a key role in *M. tuberculosis* pathogenicity, representing a promising target for tuberculosis drug intervention, and a molecular tool to dissect the unique properties of intracellular pathogens to parasitize host cells.

## 6. Immune response to tuberculosis

Control of *M. tuberculosis* is mainly the result of productive teamwork between macrophages and T-cells. T lymphocytes are central to the control of *M. tuberculosis* infection (Fig.7). The pathogen resides in the phagosome of macrophages, where mycobacterial peptides have access to the major histocompatibility complex class II (MHCII) molecules, that are shuttled to the cell surface and stimulate CD4 T cells. MHCI-restricted CD8 T cells are also stimulated by mycobacterial peptides, although the underlying mechanisms remain to be understood. It is thought that the major pathway responsible for stimulation of MHCI restricted T cell is the cross-priming (Kaufmann, 2004). *M. tuberculosis* induces apoptosis in infected macrophages and this results in the formation of apoptotic vesicles carrying mycobacterial peptides. These vesicles are taken up by dendridic cells (DC) and peptides are presented by MHCI molecules. Another hypothesis, which still remains controversial though supported by solid evidences, is the escape of mycobacterial bacilli into the cytosol. In 2007, van der Wel *et al.* published cryo-immunogold TEM images of Mtb-infected human dendritic cells and macrophages, in which Mtb was found in increasing numbers in the cytosol two days after infection (van der Wel *et al.*, 2007). Recently, Mtb has been shown to permealize phagosomal membranes in an ESX-1-dependent manner when infecting macrophages (Simeone *et al.*, 2012).

Another class of antigen-presentation molecules involved in immune response in TB is the CD1 molecules (Schaible and Kaufmann, 2000). It seems that mycobacterial glycolipids separate from mycobacteria and are incorporated into vesicles which are shuttled throughout the cell and their content is presented by CD1 molecules to

various CD1-restricted T lymphocytes. T cells that express γδ TCR also participate in the immune response against *M. tuberculosis* (Kaufmann, 1996). Human γδ T cells are stimulated by a unique group of non-proteinaceous antigens containing phosphate, named phospholigands. Mycobacterial phospholigands, such as prenyl pyrophosphates, reach the cell surface through an unknown pathway and are recognized by γδ T cells. These four cell types act in cooperation to overcome the bacterial persistence. CD4 T cells are polarized into T helper cells Th1 and Th17, which perform effector functions producing multiple cytokines, notably IL-2 for T-cell activation, IFN-γ and TNF-α for macrophage activation (Kaufmann, 2010). CD8 T cells produce IFN-γ and TNF-α to activate macrophages and they also act as cytolytic T lymphocytes (CTL) by secreting perforin and granulysin, which lyse host cells and directly attack *M. tuberculosis* (Kaufmann, 2010). These effector T cells are succeeded by memory T cells which enhance the immune response and establish its progression over the time. The large γδ T cell population readily produces IFN-γ after stimulation with phospholigands and expresses granule-dependent mycobacteriocidal activity. This cell type is therefore responsible for mobilizing the first line of defense against tuberculosis (Kaufmann, 2001; Raja, 2004). Although CD4 T cells, CD8 T cells, γδ T cells and CD1 restricted T cells are apparently all required for optimum protection, a hierarchy seems to exist: CD4 T cells are the most important players, followed by CD8 T cells.

T cells are the major mediators of protection, but macrophages provide the real effector function. The cross talk between T cells and macrophages is achieved by various cytokines, such as IFN-γ and TNF-α. The results of this cross talk are, first the macrophage activation, which means the ability of macrophages to control mycobacterial multiplication, for instance by generation of reactive oxygen and nitrogen intermediates (ROI and RNI). Secondly, cross talk is essential for establishment of well organized granulomas, where macrophages, dendritic cells, and different T cell populations exist in near vicinity. As long as this cross talk is well balanced, productive granulomas develop, resulting in containment of bacterial mass. Once any element of this complex interplay is impaired, the balance is tipped and the productive granuloma can no longer be sustained. In this case, bacteria

overcome the immune barrier and start spreading out of the granuloma. Thus, the active disease develops (Kaufmann, 2002).

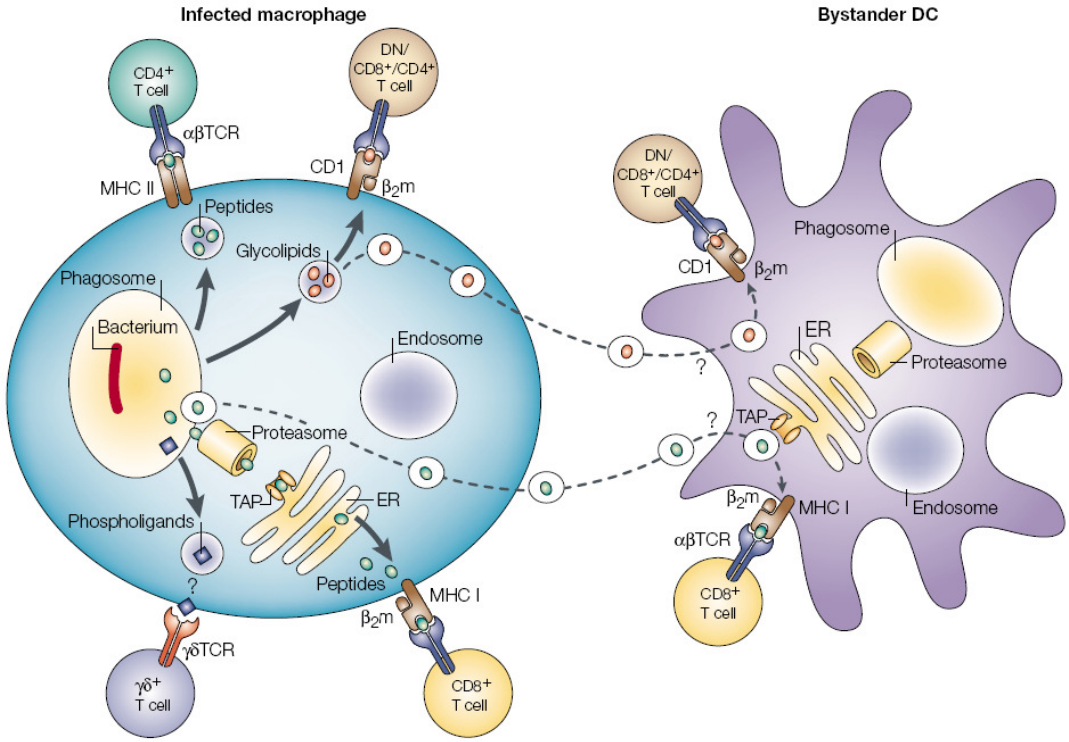



Fig.7 The immune response to *M. tuberculosis* infection (Kaufmann, 2001).

## 7. Vaccination strategies against tuberculosis - Present candidates

Although the host immune defense is onset upon infection, *M. tuberculosis* is able to evade very successfully the immune mechanisms. The pathogen is equipped with numerous immune evasion strategies, including mycobacterial secreted proteins such as superoxide dismutase, which are antagonistic to ROI. The suppression of antigen presentation processes is by far the most sophisticated strategy. *M. tuberculosis* is able to decrease MHCII molecule synthesis by inhibiting the transcriptional regulator of this class of molecules, named CIITA (Harding and Boom, 2010). Cross-priming is also reduced in Mtb infection because of mycobacterial gene products which inhibit host cell apoptosis, such as *nuoG* (Velmurugan et al., 2007). The phagosomal localization of Mtb partially protects it

from MHCI antigen presentation. Understanding the mechanisms behind the evasion of host immune response by Mtb is the basis for rational vaccine design.

The search for a vaccine against tuberculosis began 110 years ago. Robert Koch, eight years after his discovery of the tubercle bacillus (1882), devised a therapeutic vaccine termed tuberculin (semipurified supernatants of Mtb cultures) which completely failed in a large-scale clinical trial (Kaufmann and Gengenbacher, 2012). A decade later, the french microbiologists Albert Calmette and Camille Guérin performed a second attempt. After more than 200 passages, they obtained an attenuated strain of *M. bovis*, the aetiological agent of bovine tuberculosis. The two scientists proved safety and protection for the newly derived strain, upon intensive characterization in various experimental animal models (Kaufmann and Gengenbacher, 2012). The strain, named BCG (Bacillus Calmette-Guérin), is nowadays the only approved vaccine against TB. Almost three billion doses have been used since 1921. Approximately 115 million doses are distributed each year, covering 80% of infant population worldwide (Skeiky and Sadoff, 2006). BCG protects against severe forms of tuberculosis in newborn babies, as originally proposed by its discoverers (Kaufmann *et al.*, 2010). However, the protection induced by BCG vaccination against adult tuberculosis is insufficient. The lack of effectiveness in adults might be explained considering different factors. 1) A previous exposure to environmental mycobacteria, very likely in endemic areas like India, interferes with the immune response to BCG. 2) BCG has become too attenuated through culture, and modern preparations of the vaccine are too benign to generate a sufficient immunity. 3) Clearance of BCG in some individuals may occur before development of protective immune response. 4) Failure of BCG to stimulate adequate anti-mycobacterial CD4 and CD8 T cell responses (Russell *et al.*, 2010). Nevertheless, BCG is still the only vaccine in use for control of the global burden of tuberculosis. BCG represents a starting point on which TB research is working for years with the aim to improve its efficacy and overcome its weaknesses.

New anti-TB vaccination strategies can be divided into three categories. First, improving BCG by adding or overexpressing Mtb antigens, which would enhance the immune response induced by the recombinant bacterium. Second, creating attenuated Mtb strains by deletion of genes involved in survival or virulence of Mtb.

Third, the use of prime-boost strategies that amplify an initial protective immune response (conferred by live vaccine) by subsequent inoculation of Mtb antigens or protein subunits. Current pre-exposure vaccination strategies aim at reducing the initial *M. tuberculosis* bacterial burden and preventing reactivation of latent infection. Post-exposure vaccines are needed to target dormant bacteria and prevent their reactivation. They would ideally also prevent reinfection of individuals living in endemic areas.

A total of 12 vaccine candidates have entered clinical trials within the last years (Table 1). Clinical trials for TB vaccines are divided in three phases (Kaufmann, 2010). Phase I trials are safety and immunogenicity testing performed in small groups (10 or more individuals/group), typically in the region where the vaccine candidate was developed, and then repeated in highly endemic area. Phase II trials involve larger groups (100 or more individuals/group), for assessment of route of administration and optimal dosage. Finally, Phase III trials involve 20,000-50,000 participants, lasting several years. They aim to assess vaccine efficacy against natural infection in highly endemic area. Four vaccine candidates have entered or passed Phase II clinical trials. Although there is still no certainty that they will confer robust protection against TB in humans, some look promising. Most are pre-exposure vaccines and will most likely prevent tuberculosis disease. They can be divided in three groups: live mycobacteria vaccines to replace BCG, subunits vaccines for booster of BCG, and killed mycobacteria vaccines as adjunct.

**Live mycobacteria vaccines to replace BCG**

Live mycobacteria vaccines are based either on improvement of BCG vaccine by addition of relevant genes or on attenuation of *M. tuberculosis* by deletion of virulence genes. Three recombinant BCG (rBCG) vaccines have already entered initial clinical trials, after assessment of higher potency and safety in preclinical studies, compared to the current BCG vaccine. These are rBCG30, rBCG *ureC::Hly* (VPM1002) and rBCG *ureC::pfoA Rv3407*$^{+}$*fbpB*$^{+}$*fbpA*$^{+}$ (AERAS-422). The rBCG30 overexpresses Ag85B, which is shared by BCG and Mtb, and it is the first candidate to show better protection in different animal models than the parental BCG (Tullius *et al.*, 2008).

## SECTION I: Candidates Tested in Clinical Trials

| Status | Products | Product Description | Sponsors | Indication | Type of Vaccine | Target Populations |
|---|---|---|---|---|---|---|
| Phase III | Mw [M. indicus pranii (MIP)] | Whole cell saprophytic non - TB mycobacterium | Department of Biotechonology (Ministry of Science & Technology, Government of India), M/s. Cadila Pharmaceuticals Ltd. | IT | Whole cell, Inactivated or Disrupted | – |
| Phase IIb | MVA85A/ AERAS-485 | Modified vaccinia Ankara vector expressing Mtb antigen 85A | Oxford-Emergent Tuberculosis Consortium (OETC), Aeras | B, PI, IT | Viral Vectored | BCG-vaccinated infants and adolescents; HIVinfected adults |
| | AERAS-402/ Crucell Ad35 | Replication - deficient adenovirus 35 vector expressing Mtb antigens 85A, 85B, TB10.4 | Crucell, Aeras | B | Viral Vectored | BCG-vaccinated infants, children and adults |
| Phase II | M72 + AS01 | Recombinant protein composed of a fusion of Mtb antigens Rv1196 and Rv0125 & adjuvant AS01 | GSK, Aeras | B, PI | Recombinant Protein | Adolescents/ adults, infants |
| | Hybrid-I+IC31 | Adjuvanted recombinant protein composed of Mtb antigens 85B and ESAT-6 | Statens Serum Institute (SSI), TBVI, EDCTP, Intercell | P, B, PI | Recombinant Protein | Adolescents; adults |
| | VPM 1002 | rBCG Prague strain expressing listeriolysin and carries a urease deletion mutation | Max Planck, Vakzine Projekt Management GmbH, TBVI | P, B | Recombinant Live | – |
| | RUTI | Fragmented Mtb cells | Archivel Farma, S.L.. | B, PI, IT | Whole cell, Inactivated or Disrupted | HIV+ adults, LTBI diagnosed |
| Phase I | AdAg85A | Replication - deficient adenovirus 5 vector expressing Mtb antigen 85A | McMaster University | P, B, PI | Viral Vectored | Infants; adolescents; HIV+ |
| | Hybrid-I+CAF01 | Adjuvanted recombinant protein composed of Mtb antigens 85B and ESAT-6 | SSI, TBVI | P, B, IT | Recombinant Protein | Adolescents, adults |
| | Hybrid 56 + IC31 | Adjuvanted recombinant protein composed of Mtb antigens 85B, ESAT-6 and Rv2660 | SSI, Aeras, Intercell | P, B, PI | Recombinant Protein | Adolescents, adults |
| | HyVac 4/ AERAS-404, + IC31 | Adjuvanted recombinant protein composed of a fusion of Mtb antigens 85B and TB10.4 | SSI, sanofi-pasteur, Aeras, Intercell | B | Recombinant Protein | Infants |
| | AERAS-422 | Recombinant BCG expressing mutated PfoA and overexpressing antigens 85A, 85B, and Rv3407 | Aeras | P | Recombinant Live | Infants |
| Phase III [completed] | M. vaccae | Inactivated whole cell non - TB mycobacterium; phase III in BCG - primed HIV+ population completed; reformulation pending | NIH, Immodulon | B, PI, IT | Whole cell, Inactivated or Disrupted | BCG-vaccinated HIV+ adults |
| Phase I [completed] | rBCG30 | rBCG Tice strain expressing 30 kDa Mtb antigen 85B | UCLA, NIH, NIAID, Aeras | B, PI | Recombinant Live | Newborns, adolescents, and adults |
| | M. smegmatis | Whole cell extract | – | B, PI, IT | Whole cell, Inactivated or Disrupted | – |

## SECTION II: Candidates in Preclinical Studies - 2011

| Type of Vaccine | Products | Product Description | Sponsor | Indication |
|---|---|---|---|---|
| Recombinant Live | Mtb [Δ*leuCD* Δ*panCD* Δ*secA2*] | Non - replicating, Mtb strain auxotrophic for leucine and pantothenate; attenuated for secA2 | Albert Einstein College of Medicine | P |
| | MTBVAC [Δ*phoP*, Δ*fadD26*] | Live vaccine based on attenuation of Mtb by stable inactivation by deletion of phoP and fadD26 genes without antibiotic resistance markers in compliance with 2005 and 2010 Geneva consensus safety requirements | University of Zaragoza, Institute Pasteur, BIOFABRI, TBVI | P |
| Protein | HBHA | Naturally methylated 21 - kDa purified protein from M.bovis BCG | Institute Pasteur of Lille, INSERM, TBVI, Aeras | P, B, PI, IT |
| DNA | HG85A | DNA vaccines—Ag85A | Shanghai H&G Biotech | B, IT |
| | Hsp DNA vaccine | Codon - optimized heat shock protein from M. leprae, a CpG island | Sequella, Shanghai Public Health Clinical Center | B, PI, IT |

Table 1. In Section I are described the candidates currently tested in clinical trials. In the Section II are listed the candidates involved in preclinical studies. P: Prime; B: Boost; PI: Candidate is indicated post-infection; IT: Candidate is indicated for immunotherapy (Source: Stop TB Partnership, 2011).

This vaccine has successfully completed Trial I, showing a comparable safety profile to the parental BCG strain and a proof-of-principle that recombinant BCGs overexpressing immunodominant Mtb antigens are a promising approach (Rowland and McShane, 2011).The rBCG *ureC::Hly* is a recombinant strain that secretes lysteriolysin produced by *Listeria monocytogenes* and carries a deletion of *ureC* gene (Grode *et al.*, 2005). Listeriolysin is a pore-forming toxin protein which exerts its activity only at acidic pH. To maintain acidic pH in BCG phagosome, the urease gene (*ureC*), contributing for pH neutralization of the phagosome, was deleted. The recombinant bacterium therefore secretes active lysteriolysin, which by punching holes in the phagosome membrane, determines two consequences. First, perforation of the membrane facilitates translocation of mycobacterial antigens into the cytosol and subsequent MHCI loading. Second, membrane perforation promotes leakage of phagolysosomal proteases such as cathepsins into the cytosol, activating caspase cascade-inducing apoptotic cell death. Apoptosis in turn results in cross-priming of mycobacterial antigens. The potent MHCI presentation, enhancement of apoptosis, increased cross-priming and uptake by dendridic cells, result in increased CD8 and CD4 T-cell responses as well as γδ T-cell and CD1-restricted T cell activation. This vaccine is thought to overcome some weaknesses of BCG vaccine such as ineffective CD8 T cell stimulation and poor apoptosis induction in infected macrophages. In preclinical mouse-protection studies, rBCG *ureC::Hly* induced superior protective efficacy than the parental strain against Mtb challenge, along with higher safety profile in immuno-compromised SCID mice. Moreover, unlike the parental BCG strain, rBCG *ureC::Hly* was demonstrated to protect significantly against the virulent *M. tuberculosis* Beijing/W strain (Grode *et al.*, 2005). A Phase I clinical trial evaluating the safety and immunogenicity of this vaccine in healthy male volunteers has been completed and a dose-escalation RCT (randomized controlled trial) in South Africa is ongoing (Rowland and McShane, 2011).

The rBCG *ureC::pfoA Rv3407⁺fbpB⁺fbpA⁺* combines concepts underlying VPM1002 and rBCG30 (Sun *et al.*, 2009). It overexpresses the immunodominant antigens Ag85A (*fbpA*), Ag85B (*fbpB*) and Rv3407, which has been associated with latent TB. Here the recombinant bacterium expresses another type of pore-forming component, which is PFO, the perforin of the extracellular bacterium *Clostridium perfringens.* Shrewdly, PFO has been engineered by a single amino acid exchange to maintain hemolytic activity at pH 7.0. This strain combines the increased protection seen with rBCG30 with the increased protection and safety of the endosome-escape mutants. A Phase I clinical trial to evaluate safety in adults had to be terminated because two vaccinated trial participants developed shingles, probably due to reactivation of dormant herpes varicella–zoster virus. A careful elucidation of underlying mechanisms of this vaccine is therefore required (Kaufmann and Gengenbacher, 2012).

Vaccines based on *M. tuberculosis* attenuation have to carry at least two independent deletions in order to avoid spontaneous conversion to virulence. The most promising candidates are Mtb ΔRD1Δ*panCD* and Mtb Δ*phoP*Δ*fad*. The first is devoid of the entire Mtb virulence gene island RD1 (region of difference one) and auxotrophic because of pantothenate deficiency (Sambandamurthy *et al.*, 2006). The RD1 genomic region is present in all virulent strains of *M. tuberculosis*, and missing from the vaccine strain *M. bovis* BCG (Ganguly *et al.*, 2008). RD1 has been considered crucial in the pathogenesis of *M. tuberculosis.* It encodes part of the ESX-1 secretion system and some of its substrates, e.g. the two secretory proteins CFP-10 and ESAT-6 (Guinn *et al.*, 2004). The ΔRD1Δ*panCD* double deletion mutant of *M. tuberculosis* H37Rv was more attenuated and efficacious as a vaccine than was BCG, when assessed in immunocompromised mice (Sambandamurthy *et al.*, 2006). In the mutant Mtb Δ*phoP*Δ*fad*, inactivation of PhoP, which regulates expression of many Mtb proteins, along with the deletion of *fad* gene, essential for the phthiocerol dimycocerosate synthesis, attenuate *M. tuberculosis* virulence while maintaining the immune response normally associated with the parental strain (Martin *et al.*, 2006). These vaccines are expected to enter clinical trials in the next few years (Kaufmann *et al.*, 2010).

In order to improve efficacy of BCG replacement candidates, some strategies have already been planned for the near future. Introduction of genes encoding dormancy

antigens such as the DosR-regulated gene products might be an approach for post-exposure vaccination of latently infected individuals. Expression of cytokines genes such as IL-18 and IFN-γ could enhance Th1 cell polarization. Deletion of anti-apoptotic genes is a potential powerful strategy for facilitating cross-priming. A promising vaccine candidate for long-term protection is the recently generated rBCG *ureC::Hly* expressing latency-associated antigens, such as Rv2659c, Rv3407 and Rv1733c. The expression of these three latency-associated antigens significantly improved protection in mice after intra-dermal vaccination at day 200 p.i. (Reece *et al.*, 2011).

**Subunits vaccines for booster of BCG**

Seven candidates are subunit booster vaccines to be given on the top of the BCG prime. Subunit vaccine candidates are based on antigens that activate T cells in previously vaccinated humans and animals. Two types of product have been developed: recombinant fusion proteins consisting of two or three dominant Mtb or BCG antigens and live viral vectors expressing one or several mycobacterial proteins. The limiting factors of these approaches are that the first type needs an adjuvant to promote Th1 immune response, whereas the second type can be inhibited by previous exposure to the vector. Subunits vaccines trigger a Th1-dominated immune response to the expressed heterologous antigen and induce CD8 T-cell response (Kaufmann *et al.*, 2010).

Ag85B-ESAT-6 (H1) candidate vaccine was designed with a strong Th1 adjuvant, IC31. This adjuvant is a mixture of oligodeoxynucleotides, that stimulate dendridic cells, and polycationic amino acids, that increase antigen uptake by APC (Antigen-Presenting Cell). The vaccine has been assessed in two Phase I clinical trials, demonstrating high tolerability and immunogenicity in humans. In fact, it was shown to induce a strong persisting Th1 response to the Mtb Ag85 protein, without interfering with ESAT-6-based diagnostic assays performed a few months after immunization. The same fusion protein was also used for formulation with a novel liposomal adjuvant, CAF01, giving rise to durable, multifunctional Th1 T-cell responses and protection against Mtb aerosol challenge in mice (Rowland and McShane, 2011). An alternative fusion protein was developed, replacing ESAT-6 with

TB-10.4 (*esxH*), to reduce the putative risk of interference with diagnostic tools. The vaccine was named Ag85B-TB-10.4 (AERAS-404).

M72 is a recombinant fusion of Mtb39 and Mtb32 proteins, formulated in the vaccine adjuvant system AS02A, which is an immuno-potentiator composed of monophosphoryl lipid A, an immuno-stimulant, and the detergent QS21. The two fusion proteins were selected for their ability to induce IFN- γ production by both CD4 and CD8 T-cells. The vaccine completed Phase I and II, showing acceptable tolerability and profound T-cell responses in healthy adults (Von Eschen *et al.*, 2009).

MVA85A is a recombinant, replication-deficient vaccinia virus, expressing Ag85A. The promising safety and immunogenicity of this candidate vaccine that arose from the Phase I clinical trials in UK led to further Phase I and IIa clinical trials in target populations in South Africa, The Gambia and Senegal (Rowland and McShane, 2011; Sander *et al.*, 2009).

Ad35 (AERAS-402) and AdAg85A are replication-deficient recombinant adenoviruses (Ad35 and Ad5, respectively) expressing Ag85A, Ag85B and TB10.4 and solely Ag85A, respectively (Radosevic *et al.*, 2007; Santosuosso *et al.*, 2006). These vaccines are of special interest for their excellent ability to stimulate CD8 T-cells. The counterpart is the diminished effectiveness due to a previous exposure to a homologous or cross-reacting strain of adenovirus. This event is quite likely, since Adenovirus 5 is highly prevalent in the human population, while Adenovirus 35 has a fairly low prevalence (ranging from 5% in developed countries to 20% in Africa) (Kaufmann *et al.*, 2010).

**Killed mycobacteria vaccines as adjunct**

The rationale of the therapy based on killed mycobacteria vaccines is first to take advantage of the bactericidal properties of chemotherapy to kill active growing bacilli and reduce local inflammatory responses. After chemotherapy, killed mycobacteria can be inoculated to reduce the probability of re-growth of the remaining latent bacilli.

Whole heat-killed *M. vaccae*, an environmental saprophyte, has been developed as an immuno-therapeutic agent against tuberculosis (Dlugovitzky *et al.*, 2006). The vaccine is usually given intradermally in a multidose series. Variable results have been obtained in different geographical locations within African continent. Clinical

studies show that it might improve radiological and clinical outcomes, but the discrepant results led the vaccine to be evaluated as a prophylactic vaccine (Kaufmann *et al.*, 2010) .

RUTI is another proposed therapeutic vaccine candidate. *M. tuberculosis* is grown under stress conditions and bacilli are then fragmented, detoxified, and delivered in liposomes (Cardona, 2006). In a Phase I clinical trial in Spain it showed promising results, but it still has to be evaluated in TB positive individuals (Rowland and McShane, 2011).

Three types of vaccines are ideally to be considered in a comprehensive tuberculosis vaccination strategy (Kaufmann *et al.*, 2010). First, an improved BCG vaccine should prime the immune system, meaning stimulation of T cells that are subsequently ready to be re-stimulated by boost vaccination. In case of early tuberculosis exposure, such a vaccine would at least slow down bacterial growth, reduce bacterial burden and therefore protect against tuberculosis disease. After a few weeks to several years, a subunits vaccine should be given. The aim would be to boost and re-orient the BCG-induced memory response. Such boosting might have to be repeated during an individual's lifespan, with changing antigen composition. To target latent infection and prevent reactivation of dormant Mtb, a post-exposure vaccine should be given to adolescent and adults carrying a latent TB infection. This last category of vaccines are less well advanced than the pre-exposure ones, and the present state-of-art of post-exposure vaccines entering clinical trials in 2010 includes one candidate: Hybrid 56, a formulation of the Ag85B-ESAT-6 with the addition of the dormancy antigen Rv2660 in IC31 (Kaufmann *et al.*, 2010).

## REFERENCES


Abdallah, A.M., Gey van Pittius, N.C., Champion, P.A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C.M., Appelmelk, B.J., and Bitter, W. (2007) Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* **5**: 883-891.

Adekoya, O.A., and Sylte, I. (2009) The thermolysin family (M4) of enzymes: therapeutic and biotechnological potential. *Chem Biol Drug Des* **73**: 7-16.

Agard, N.J., and Wells, J.A. (2009) Methods for the proteomic identification of protease substrates. *Curr Opin Chem Biol* **13**: 503-509.

Algood, H.M., Chan, J., and Flynn, J.L. (2003) Chemokines and tuberculosis. *Cytokine Growth Factor Rev* **14**: 467-477.

Balcewicz-Sablinska, M.K., Keane, J., Kornfeld, H., and Remold, H.G. (1998) Pathogenic Mycobacterium tuberculosis evades apoptosis of host macrophages by release of TNF-R2, resulting in inactivation of TNF-alpha. *J Immunol* **161**: 2636-2641.

Bashir, N., Kounsar, F., Mukhopadhyay, S., and Hasnain, S.E. (2010) Mycobacterium tuberculosis conserved hypothetical protein rRv2626c modulates macrophage effector functions. *Immunology* **130**: 34-45.

Bland, N.D., Pinney, J.W., Thomas, J.E., Turner, A.J., and Isaac, R.E. (2008) Bioinformatic analysis of the neprilysin (M13) family of peptidases reveals complex evolutionary and functional relationships. *BMC Evol Biol* **8**: 16.

Booth, B.A., Boesman-Finkelstein, M., and Finkelstein, R.A. (1984) Vibrio cholerae hemagglutinin/protease nicks cholera enterotoxin. *Infect Immun* **45**: 558-560.

Bredemeyer, A.J., Townsend, R.R., and Ley, T.J. (2005) Use of protease proteomics to discover granzyme B substrates. *Immunol Res* **32**: 143-153.

Brennan, P.J., and Nikaido, H. (1995) The envelope of mycobacteria. *Annu Rev Biochem* **64**: 29-63.

Cardona, P.J. (2006) RUTI: a new chance to shorten the treatment of latent tuberculosis infection. *Tuberculosis (Edinb)* **86**: 273-289.

Clark, V.L., Bavoil, P.M. (2002) *Methods in enzymology*: Academic Press.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd, Tekaia, F., Badcock, K., Basham, D.,

Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., and Barrell, B.G. (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* **393**: 537-544.

Davis, J.M., and Ramakrishnan, L. (2009) The role of the granuloma in expansion and dissemination of early tuberculous infection. *Cell* **136**: 37-49.

Dlugovitzky, D., Fiorenza, G., Farroni, M., Bogue, C., Stanford, C., and Stanford, J. (2006) Immunological consequences of three doses of heat-killed Mycobacterium vaccae in the immunotherapy of tuberculosis. *Respir Med* **100**: 1079-1087.

Ducati, R.G., Ruffino-Netto, A., Basso, L.A., and Santos, D.S. (2006) The resumption of consumption -- a review on tuberculosis. *Mem Inst Oswaldo Cruz* **101**: 697-714.

Ferraris, D.M., Sbardella, D., Petrera, A., Marini, S., Amstutz, B., Coletta, M., Sander, P., and Rizzi, M. (2011) Crystal structure of Mycobacterium tuberculosis zinc-dependent metalloprotease-1 (Zmp1), a metalloprotease involved in pathogenicity. *J Biol Chem* **286**: 32475-32482.

Fratti, R.A., Backer, J.M., Gruenberg, J., Corvera, S., and Deretic, V. (2001) Role of phosphatidylinositol 3-kinase and Rab5 effectors in phagosomal biogenesis and mycobacterial phagosome maturation arrest. *J Cell Biol* **154**: 631-644.

Fratti, R.A., Chua, J., Vergne, I., and Deretic, V. (2003) Mycobacterium tuberculosis glycosylated phosphatidylinositol causes phagosome maturation arrest. *Proc Natl Acad Sci U S A* **100**: 5437-5442.

Frieden, T.R., Sterling, T.R., Munsiff, S.S., Watt, C.J., and Dye, C. (2003) Tuberculosis. *Lancet* **362**: 887-899.

Ganguly, N., Siddiqui, I., and Sharma, P. (2008) Role of M. tuberculosis RD-1 region encoded secretory proteins in protective response and virulence. *Tuberculosis (Edinb)* **88**: 510-517.

Grode, L., Seiler, P., Baumann, S., Hess, J., Brinkmann, V., Nasser Eddine, A., Mann, P., Goosmann, C., Bandermann, S., Smith, D., Bancroft, G.J., Reyrat, J.M., van Soolingen, D., Raupach, B., and Kaufmann, S.H. (2005) Increased vaccine efficacy against tuberculosis of recombinant Mycobacterium bovis bacille Calmette-Guerin mutants that secrete listeriolysin. *J Clin Invest* **115**: 2472-2479.

Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., Smith, S., and Sherman, D.R. (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis. *Mol Microbiol* **51**: 359-370.

Harding, C.V., and Boom, W.H. (2010) Regulation of antigen presentation by Mycobacterium tuberculosis: a role for Toll-like receptors. *Nat Rev Microbiol* **8**: 296-307.

Hernick, M., and Fierke, C.A. (2005) Zinc hydrolases: the mechanisms of zinc-dependent deacetylases. *Arch Biochem Biophys* **433**: 71-84.

Hwang, I.K., Park, S.M., Kim, S.Y., and Lee, S.T. (2004) A proteomic approach to identify substrates of matrix metalloproteinase-14 in human plasma. *Biochim Biophys Acta* **1702**: 79-87.

Ichinose, Y., Ehara, M., Honda, T., and Miwatani, T. (1994) The effect on enterotoxicity of protease purified from Vibrio cholerae O1. *FEMS Microbiol Lett* **115**: 265-271.

Indrigo, J., Hunter, R.L., Jr., and Actor, J.K. (2003) Cord factor trehalose 6,6'-dimycolate (TDM) mediates trafficking events during mycobacterial infection of murine macrophages. *Microbiology* **149**: 2049-2059.

Issaq, H., and Veenstra, T. (2008) Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): advances and perspectives. *Biotechniques* **44**: 697-698, 700.

Iwata, N., Tsubuki, S., Takaki, Y., Watanabe, K., Sekiguchi, M., Hosoki, E., Kawashima-Morishima, M., Lee, H.J., Hama, E., Sekine-Aizawa, Y., and Saido, T.C. (2000) Identification of the major Abeta1-42-degrading catabolic pathway in brain parenchyma: suppression leads to biochemical and pathological deposition. *Nat Med* **6**: 143-150.

Johansen, P., Fettelschoss, A., Amstutz, B., Selchow, P., Waeckerle-Men, Y., Keller, P., Deretic, V., Held, L., Kundig, T.M., Bottger, E.C., and Sander, P. (2011) Relief from Zmp1-mediated arrest of phagosome maturation is associated with facilitated presentation and enhanced immunogenicity of mycobacterial antigens. *Clin Vaccine Immunol* **18**: 907-913.

Kaufmann, S.H. (1996) gamma/delta and other unconventional T lymphocytes: what do they see and what do they do? *Proc Natl Acad Sci U S A* **93**: 2272-2279.

Kaufmann, S.H. (2001) How can immunology contribute to the control of tuberculosis? *Nat Rev Immunol* **1**: 20-30.

Kaufmann, S.H. (2002) Protection against tuberculosis: cytokines, T cells, and macrophages. *Ann Rheum Dis* **61 Suppl 2**: ii54-58.

Kaufmann, S.H. (2004) New issues in tuberculosis. *Ann Rheum Dis* **63 Suppl 2**: ii50-ii56.

Kaufmann, S.H., Cole, S.T., Mizrahi, V., Rubin, E., and Nathan, C. (2005) Mycobacterium tuberculosis and the host response. *J Exp Med* **201**: 1693-1697.

Kaufmann, S.H. (2010) Future vaccination strategies against tuberculosis: thinking outside the box. *Immunity* **33**: 567-577.

Kaufmann, S.H., Hussey, G., and Lambert, P.H. (2010) New vaccines for tuberculosis. *Lancet* **375**: 2110-2119.

Kaufmann, S.H., and Gengenbacher, M. (2012) Recombinant live vaccine candidates against tuberculosis. *Curr Opin Biotechnol*.

Keller, P.M., Bottger, E.C., and Sander, P. (2008) Tuberculosis vaccine strain Mycobacterium bovis BCG Russia is a natural recA mutant. *BMC Microbiol* **8**: 120.

Koul, A., Herget, T., Klebl, B., and Ullrich, A. (2004) Interplay between mycobacteria and host signalling pathways. *Nat Rev Microbiol* **2**: 189-202.

Lawn, S.D., Butera, S.T., and Shinnick, T.M. (2002) Tuberculosis unleashed: the impact of human immunodeficiency virus infection on the host granulomatous response to Mycobacterium tuberculosis. *Microbes Infect* **4**: 635-646.

Lawn, S.D., and Zumla, A.I. (2011) Tuberculosis. *Lancet* **378**: 57-72.

Lee, A.Y., Goo Park, S., Kho, C.W., Young Park, S., Cho, S., Lee, S.C., Lee, D.H., Myung, P.K., and Park, B.C. (2004) Identification of the degradome of Isp-1, a major intracellular serine protease of Bacillus subtilis, by two-dimensional gel electrophoresis and matrix- assisted laser desorption/ionization-time of flight analysis. *Proteomics* **4**: 3437-3445.

Lopez-Otin, C., and Overall, C.M. (2002) Protease degradomics: a new challenge for proteomics. *Nat Rev Mol Cell Biol* **3**: 509-519.

MacMicking, J.D., Taylor, G.A., and McKinney, J.D. (2003) Immune control of tuberculosis by IFN-gamma-inducible LRG-47. *Science* **302**: 654-659.

Maiti, D., Bhattacharyya, A., and Basu, J. (2001) Lipoarabinomannan from Mycobacterium tuberculosis promotes macrophage survival by phosphorylating Bad through a phosphatidylinositol 3-kinase/Akt pathway. *J Biol Chem* **276**: 329-333.

Makinoshima, H., and Glickman, M.S. (2005) Regulation of Mycobacterium tuberculosis cell envelope composition and virulence by intramembrane proteolysis. *Nature* **436**: 406-409.

Martin, C., Williams, A., Hernandez-Pando, R., Cardona, P.J., Gormley, E., Bordat, Y., Soto, C.Y., Clark, S.O., Hatch, G.J., Aguilar, D., Ausina, V., and Gicquel, B. (2006) The live Mycobacterium tuberculosis phoP mutant strain is more attenuated than BCG and confers protective immunity against tuberculosis in mice and guinea pigs. *Vaccine* **24**: 3408-3419.

Master, S.S., Rampini, S.K., Davis, A.S., Keller, C., Ehlers, S., Springer, B., Timmins, G.S., Sander, P., and Deretic, V. (2008) Mycobacterium tuberculosis prevents inflammasome activation. *Cell Host Microbe* **3**: 224-232.

Miyoshi, S., Hirata, Y., Tomochika, K., and Shinoda, S. (1994) Vibrio vulnificus may produce a metalloprotease causing an edematous skin lesion in vivo. *FEMS Microbiol Lett* **121**: 321-325.

Miyoshi, S., and Shinoda, S. (2000) Microbial metalloproteases and pathogenesis. *Microbes Infect* **2**: 91-98.

Oefner, C., D'Arcy, A., Hennig, M., Winkler, F.K., and Dale, G.E. (2000) Structure of human neutral endopeptidase (Neprilysin) complexed with phosphoramidon. *J Mol Biol* **296**: 341-349.

Panchal, R.G., Hermone, A.R., Nguyen, T.L., Wong, T.Y., Schwarzenbacher, R., Schmidt, J., Lane, D., McGrath, C., Turk, B.E., Burnett, J., Aman, M.J., Little, S., Sausville, E.A., Zaharevitz, D.W., Cantley, L.C., Liddington, R.C., Gussio, R., and Bavari, S. (2004) Identification of small molecule inhibitors of anthrax lethal factor. *Nat Struct Mol Biol* **11**: 67-72.

Pelicic, V., Reyrat, J.M., and Gicquel, B. (1996) Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Mol Microbiol* **20**: 919-925.

Pethe, K., Swenson, D.L., Alonso, S., Anderson, J., Wang, C., and Russell, D.G. (2004) Isolation of Mycobacterium tuberculosis mutants defective in the arrest of phagosome maturation. *Proc Natl Acad Sci U S A* **101**: 13642-13647.

Radosevic, K., Wieland, C.W., Rodriguez, A., Weverling, G.J., Mintardjo, R., Gillissen, G., Vogels, R., Skeiky, Y.A., Hone, D.M., Sadoff, J.C., van der Poll, T., Havenga, M., and Goudsmit, J. (2007) Protective immune responses to a recombinant adenovirus type 35 tuberculosis vaccine in two mouse strains: CD4 and CD8 T-cell epitope mapping and role of gamma interferon. *Infect Immun* **75**: 4105-4115.

Raja, A. (2004) Immunology of tuberculosis. *Indian J Med Res* **120**: 213-232.

Reece, S.T., Nasser-Eddine, A., Dietrich, J., Stein, M., Zedler, U., Schommer-Leitner, S., Ottenhoff, T.H., Andersen, P., and Kaufmann, S.H. (2011) Improved long-term protection against Mycobacterium tuberculosis Beijing/W in mice after intra-dermal inoculation of recombinant BCG expressing latency associated antigens. *Vaccine* **29**: 8740-8744.

Reyrat, J.M., Lopez-Ramirez, G., Ofredo, C., Gicquel, B., and Winter, N. (1996) Urease activity does not contribute dramatically to persistence of Mycobacterium bovis bacillus Calmette-Guerin. *Infect Immun* **64**: 3934-3936.

Rojas, M., Garcia, L.F., Nigou, J., Puzo, G., and Olivier, M. (2000) Mannosylated lipoarabinomannan antagonizes Mycobacterium tuberculosis-induced macrophage apoptosis by altering Ca+2-dependent cell signaling. *J Infect Dis* **182**: 240-251.

Rowland, R., and McShane, H. (2011) Tuberculosis vaccines in clinical trials. *Expert Rev Vaccines* **10**: 645-658.

Russell, D.G. (2001) Mycobacterium tuberculosis: here today, and here tomorrow. *Nat Rev Mol Cell Biol* **2**: 569-577.

Russell, D.G. (2007) Who puts the tubercle in tuberculosis? *Nat Rev Microbiol* **5**: 39-47.

Russell, D.G., Barry, C.E., 3rd, and Flynn, J.L. (2010) Tuberculosis: what we don't know can, and does, hurt us. *Science* **328**: 852-856.

Sambandamurthy, V.K., Derrick, S.C., Hsu, T., Chen, B., Larsen, M.H., Jalapathy, K.V., Chen, M., Kim, J., Porcelli, S.A., Chan, J., Morris, S.L., and Jacobs, W.R., Jr. (2006) Mycobacterium tuberculosis DeltaRD1 DeltapanCD: a safe and limited replicating mutant strain that protects immunocompetent and immunocompromised mice against experimental tuberculosis. *Vaccine* **24**: 6309-6320.

Sander, C.R., Pathan, A.A., Beveridge, N.E., Poulton, I., Minassian, A., Alder, N., Van Wijgerden, J., Hill, A.V., Gleeson, F.V., Davies, R.J., Pasvol, G., and McShane, H. (2009) Safety and immunogenicity of a new tuberculosis vaccine, MVA85A, in Mycobacterium tuberculosis-infected individuals. *Am J Respir Crit Care Med* **179**: 724-733.

Santosuosso, M., McCormick, S., Zhang, X., Zganiacz, A., and Xing, Z. (2006) Intranasal boosting with an adenovirus-vectored vaccine markedly enhances protection by parenteral Mycobacterium bovis BCG immunization against pulmonary tuberculosis. *Infect Immun* **74**: 4634-4643.

Schaible, U.E., Sturgill-Koszycki, S., Schlesinger, P.H., and Russell, D.G. (1998) Cytokine activation leads to acidification and increases maturation of Mycobacterium avium-containing phagosomes in murine macrophages. *J Immunol* **160**: 1290-1296.

Schaible, U.E., and Kaufmann, S.H. (2000) CD1 and CD1-restricted T cells in infections with intracellular bacteria. *Trends Microbiol* **8**: 419-425.

Schulz, H., Dale, G.E., Karimi-Nejad, Y., and Oefner, C. (2009) Structure of human endothelin-converting enzyme I complexed with phosphoramidon. *J Mol Biol* **385**: 178-187.

Sharpe, M.L., Gao, C., Kendall, S.L., Baker, E.N., and Lott, J.S. (2008) The structure and unusual protein chemistry of hypoxic response protein 1, a latency

antigen and highly expressed member of the DosR regulon in Mycobacterium tuberculosis. *J Mol Biol* **383**: 822-836.

Simeone, R., Bobard, A., Lippmann, J., Bitter, W., Majlessi, L., Brosch, R., and Enninga, J. (2012) Phagosomal rupture by Mycobacterium tuberculosis results in toxicity and host cell death. *PLoS Pathog* **8**: e1002507.

Skeiky, Y.A., and Sadoff, J.C. (2006) Advances in tuberculosis vaccine strategies. *Nat Rev Microbiol* **4**: 469-476.

Sklar, J.G., Makinoshima, H., Schneider, J.S., and Glickman, M.S. (2010) M. tuberculosis intramembrane protease Rip1 controls transcription through three anti-sigma factor substrates. *Mol Microbiol* **77**: 605-617.

Stover, C.K., de la Cruz, V.F., Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T., Bansal, G.P., Young, J.F., Lee, M.H., Hatfull, G.F., and et al. (1991) New use of BCG for recombinant vaccines. *Nature* **351**: 456-460.

Sun, R., Skeiky, Y.A., Izzo, A., Dheenadhayalan, V., Imam, Z., Penn, E., Stagliano, K., Haddock, S., Mueller, S., Fulkerson, J., Scanga, C., Grover, A., Derrick, S.C., Morris, S., Hone, D.M., Horwitz, M.A., Kaufmann, S.H., and Sadoff, J.C. (2009) Novel recombinant BCG expressing perfringolysin O and the over-expression of key immunodominant antigens; pre-clinical characterization, safety and protection against challenge with Mycobacterium tuberculosis. *Vaccine* **27**: 4412-4423.

Tam, L.T., Engelbrecht, S., Talent, J.M., Gracy, R.W., and Erdos, E.G. (1985) The importance of disulfide bridges in human endopeptidase (enkephalinase) after proteolytic cleavage. *Biochem Biophys Res Commun* **133**: 1187-1192.

Tullius, M.V., Harth, G., Maslesa-Galic, S., Dillon, B.J., and Horwitz, M.A. (2008) A Replication-Limited Recombinant Mycobacterium bovis BCG vaccine against tuberculosis designed for human immunodeficiency virus-positive persons is safer and more efficacious than BCG. *Infect Immun* **76**: 5200-5214.

Turner, A.J., Isaac, R.E., and Coates, D. (2001) The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. *Bioessays* **23**: 261-269.

van der Wel, N., Hava, D., Houben, D., Fluitsma, D., van Zon, M., Pierson, J., Brenner, M., and Peters, P.J. (2007) M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell* **129**: 1287-1298.

Velmurugan, K., Chen, B., Miller, J.L., Azogue, S., Gurses, S., Hsu, T., Glickman, M., Jacobs, W.R., Jr., Porcelli, S.A., and Briken, V. (2007) Mycobacterium tuberculosis nuoG is a virulence gene that inhibits apoptosis of infected host cells. *PLoS Pathog* **3**: e110.

Vercellone, A., Nigou, J., and Puzo, G. (1998) Relationships between the structure and the roles of lipoarabinomannans and related glycoconjugates in tuberculosis pathogenesis. *Front Biosci* **3**: e149-163.

Vergne, I., Chua, J., and Deretic, V. (2003) Tuberculosis toxin blocking phagosome maturation inhibits a novel Ca2+/calmodulin-PI3K hVPS34 cascade. *J Exp Med* **198**: 653-659.

Vergne, I., Chua, J., Singh, S.B., and Deretic, V. (2004a) Cell biology of mycobacterium tuberculosis phagosome. *Annu Rev Cell Dev Biol* **20**: 367-394.

Vergne, I., Fratti, R.A., Hill, P.J., Chua, J., Belisle, J., and Deretic, V. (2004b) Mycobacterium tuberculosis phagosome maturation arrest: mycobacterial phosphatidylinositol analog phosphatidylinositol mannoside stimulates early endosomal fusion. *Mol Biol Cell* **15**: 751-760.

Vergne, I., Chua, J., Lee, H.H., Lucas, M., Belisle, J., and Deretic, V. (2005) Mechanism of phagolysosome biogenesis block by viable Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* **102**: 4033-4038.

Von Eschen, K., Morrison, R., Braun, M., Ofori-Anyinam, O., De Kock, E., Pavithran, P., Koutsoukos, M., Moris, P., Cain, D., Dubois, M.C., Cohen, J., and Ballou, W.R. (2009) The candidate tuberculosis vaccine Mtb72F/AS02A: Tolerability and immunogenicity in humans. *Hum Vaccin* **5**: 475-482.

Walburger, A., Koul, A., Ferrari, G., Nguyen, L., Prescianotto-Baschong, C., Huygen, K., Klebl, B., Thompson, C., Bacher, G., and Pieters, J. (2004) Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* **304**: 1800-1804.

# PART I – Structural and biochemical characterization of the *M. tuberculosis* zinc metallopeptidase Zmp1 and identification of potential substrates

## 1. Overview

The broad arsenal of virulence factors of *M. tuberculosis* is still poorly characterized. Recently, a new virulence determinant involved in the phagosome maturation arrest was revealed, the zinc metallopeptidase Zmp1. The peptidase was described as an important player in pathogenesis, and in the host immune response to *M. tuberculosis*. Due to its key role in mycobacterial virulence, Zmp1 represents an attractive target protein to investigate. The knowledge of this virulence factor is the prerequisite for exploiting it as putative drug target and for rationale vaccine development. The subject of the Part I is the functional and structural characterization of Zmp1. The zinc metallopeptidase Zmp1 was extensively investigated from biochemical and structural point of views, aiming to achieve the broadest knowledge about this virulence factor. The crystal structure of Zmp1 was resolved, and similarities with other eukaryotic peptidases have been revealed. In parallel, the biochemical and enzymatic properties of Zmp1 were studied. The question of which are the substrates targeted by Zmp1 was addressed. The Zmp1 substrate specificity was determined and several *in vitro* substrates have been described.

## 2. Crystal Structure of *Mycobacterium tuberculosis* Zinc-dependent Metalloprotease-1 (Zmp1), a Metalloprotease Involved in Pathogenicity

Davide M. Ferraris, Diego Sbardella, Agnese Petrera, Stefano Marini, Beat Amstutz, Massimo Coletta, Peter Sander, and Menico Rizzi

The Journal of Biological Chemistry (2011) **286**, 32475-3248

**ABSTRACT**

*Mycobacterium tuberculosis*, the causative agent of tuberculosis, parasitizes host macrophages. The resistance of the tubercle bacilli to the macrophage hostile environment relates to their ability to impair phagosome maturation and its fusion with the lysosome, thus preventing the formation of the phago-lysosome and eventually arresting the process of phagocytosis. The *M. tuberculosis* zinc-dependent metalloprotease Zmp1 has been proposed to play a key role in the process of phagosome maturation inhibition and emerged as an important player in pathogenesis. Here, we report the crystal structure of wild-type Zmp1 at 2.6 Å resolution in complex with the generic zinc metalloprotease inhibitor phosphoramidon, which we demonstrated to inhibit the enzyme potently. Our data represent the first structural characterization of a bacterial member of the zinc-dependent M13 endopeptidase family and revealed a significant degree of conservation with eukaryotic enzymes. However, structural comparison of the Zmp1-phosphoramidon complex with homologous human proteins neprilysin and endothelin-converting enzyme-1 revealed unique features of the Zmp1 active site to be exploited for the rational design of specific inhibitors that may prove useful as a pharmacological tool for better understanding Zmp1 biological function.

## INTRODUCTION

The intracellular pathogen *Mycobacterium tuberculosis* is the main etiological agent of human tuberculosis (TB), one of the world's deadliest diseases. According to the World Health Organization 2010 Fact Sheet, one third of the world population is exposed to *M. tuberculosis*, and 9 million people develop active TB, with 2 million deaths per year. The already high TB global burden is further exacerbated by the high mortality rate of TB-infected individuals suffering from HIV/AIDS. Moreover, the pharmacological treatment of TB is long and expensive, and misuse of first-line and second-line drugs against TB can favor the development of multidrug-resistant or extensively drug resistant *M. tuberculosis* strains, rendering extremely difficult the eradication of the pathogen by the most effective anti-TB drugs (2).

The host organism deals with *M. tuberculosis* infection by a series of innate and adaptive immune responses (3). Inhaled *M. tuberculosis* bacteria are phagocytosed by resident macrophages in the lungs, *i.e.* the alveolar macrophages, but not efficiently cleared (4). In immunocompetent individuals the initial acute infection is controlled by the immune system, and living bacteria are confined in a peculiar localized pulmonary structure called granuloma. There, the bacteria can endure indefinitely in a latent, nonvirulent form, and are reactivated whenever an immunosuppressive condition occurs (5).

A key step for successful bacterial clearance by macrophages is phagosome maturation process, which culminates in the formation of the phago-lysosome (6). However, some pathogens, among them *M. tuberculosis*, have evolved the ability to preclude the macrophages killing by inhibiting the phagosome maturation process (7, 8). This step is critical for the progression of the infection because it compromises bacterial clearance (9) and antigen processing (10, 11). Master *et al.* (12) proposed that *M. tuberculosis* is able to suppress phagosome maturation by inhibiting the inflammasome (13). The inflammasome is a multiprotein complex composed by members of the cytosolic sensor proteins family called nucleotide binding oligomerization domain which, once activated upon recognition of pathogen-associated molecules in the extracellular or the intracellular compartment, drives the activation of pro-caspase-1 (14, 15). Activated caspase-1, in turn, proteolitically activates pro-IL-1β into IL-1β, which, once secreted, in an autocrine and paracrine

fashion triggers the phagosome fusion with intracellular lysosomes and the early inflammatory response (16).

Master *et al.* found that *zmp1* gene (Rv0198c) suppresses inflammasome activation by inhibiting caspase-1 activation, thus preventing processing of pro-IL-1β into IL-1β and the consequent phagosome maturation. Nonetheless, they furnish evidence that suppression of the *zmp1* gene reestablished the activation of caspase-1, the production of IL-1β, and the full maturation of the phagosome into phago-lysosome, leading to the clearance of the pathogen. In addition, the authors showed that exogenously added IL-1β was sufficient to determine bacteria clearance by the phago-lysosome. The authors therefore proposed that the blocking of the phagosome maturation process is due to the inhibition of the inflammasome by the secreted *M. tuberculosis* protein Zmp1 (Zn-dependent metalloprotease-1).

In contrast to what was proposed by Master *et al.*, a recent report by Muttucumaru *et al.* (17) claims that deletion of the *zmp1* gene causes bacterial hypervirulence in a mouse model. This differs from what was observed previously by Master *et al.*, where *zmp1* deletion led to virulence attenuation. However, these two reports suggest a key role of Zmp1 during *M. tuberculosis* pathogenicity, although its mechanism of action still remains under debate.

BLAST and Pfam sequence analysis indicated that Zmp1 is an M13 endopeptidase, a protein family present in a wide range of organisms including mammals and bacteria, with the exception of yeast (18). M13 endopeptidases regulate the biological activity of many hormones and peptides and are involved in many important processes such as blood pressure regulation (neprilysin, or NEP) (19), cardiovascular development (endothelin-converting enzyme-1, or ECE-1) (20), prevention of hemolytic reaction (KELL) (21) and phosphate homeostasis (PHEX) (22). M13 endopeptidases are type II, single-pass transmembrane zinc-metallopeptidases with a hydrophobic N-terminal section of about 20 amino acids spanning the cytoplasmic membrane, and a large ectodomain of about 700 residues. Sequence analyses indicate that Zmp1, unlike NEP and ECE-1 and other members of the M13 family, lacks the N-terminal sequence required for extracellular export (although Zmp1 has been found in cell supernatants) (12) and the hydrophobic segment providing cell membrane anchoring.

All M13 endopeptidases are characterized by three signature motifs involved in the binding of $Zn^{2+}$ (HE*XX*H and E*XXX*D) and substrate/inhibitor (VNA*X*Y). The glutamate residue of the HE*XX*H signature fingerprint is essential for catalysis because it polarizes the water molecule that facilitates the nucleophilic attack to the substrate peptide bond (23).

To shed light on the so far uncharacterized process of inflammasome inhibition by *M. tuberculosis*, we solved the crystal structure of Zmp1 in complex with the inhibitor phosphoramidon. Our structural data reveal a significant structural conservation with human zinc-dependent metalloproteases NEP (24) and ECE-1 (25), thus identifying *M. tuberculosis* Zmp1 as a new member of M13 Zn-dependent metalloproteases. However, subtle differences are present in the catalytic site of Zmp1 that could be exploited for the design of specific inhibitors against the mycobacterial enzyme.

## EXPERIMENTAL PROCEDURES

*Protein Expression and Purification* - Full-length *zmp1* was cloned in pET100/d-TOPO® vector (providing an N-terminal enterokinase-cleavable His tag) and expressed in *Escherichia coli* BL21 (DE3) cells. Bacteria were precultured overnight in 2× TY medium and then diluted in 1 liter of 2× TY medium. The absorbance (*A*) was constantly monitored until it reached 0.6. The temperature was then shifted to 20 °C, and protein expression was induced overnight by the addition of 0.5 mM isopropyl 1-thio-β-d-galactopyranoside. The cells were pelleted and resuspended in 30 ml of 1× PBS, pH 7.4, and lysed using a mechanical disruption system (Basic Z apparatus; Constant System). Pellet and supernatant were separated by centrifugation, and the supernatant was applied to a preequilibrated nickel-nitrilotriacetic acid column (Qiagen) and incubated for 1 h at 4 °C. Resin was washed thoroughly with 1× PBS, pH 7.4, and protein was eluted with a step gradient of imidazole. Protein eluted in 100 and 200 mM imidazole fractions.

The eluted protein was then diluted in 10 mM Tris, pH 8, buffer and loaded on a Mono Q 5/50 column (GE Healthcare) and washed thoroughly. Protein elution was performed applying a linear gradient of NaCl. Fractions containing the purified protein were concentrated and loaded on a Sephacryl 200 16/60 gel filtration column

pre equilibrated with 10 mM Tris, pH 8. Zmp1 eluted as a monomer with a symmetric peak. Fractions containing Zmp1 were pooled and used immediately for crystallization or flash frozen in liquid nitrogen and stored at −80 °C.

*Protein Crystallization and Structure Solution* - Purified Zmp1 was concentrated to 26 mg/ml using Vivaspin concentrators (Sartorius AG) with a molecular mass cut-off of 50 kDa. Phosphoramidon ((*N*-(α-rhamnopyranosyl-oxyhydroxy-phosphinyl)-l-leucyl-l-tryptophan); catalog no. R9382, Sigma) was then added to the concentrated protein to a final concentration of 1 mM. Protein concentration of 13 mg/ml was obtained by mixing the phosphoramidon-complexed Zmp1 with gel filtration buffer in a 1:1 ratio. Initial crystallization screens were performed with an Oryx4 Protein Crystallization Robot (Douglas Instruments Ltd.). First Zmp1 crystals were obtained using the Hampton Crystal Screen 2. Optimized diffraction quality crystals were obtained by mixing 0.3 µl of reservoir solution (0.2 M ammonium sulfate, 0.1 M sodium acetate trihydrate, pH 4.6, 30% w/v polyethylene glycol monomethyl ether 2000) with 0.3 µl of phosphoramidon-complexed Zmp1 concentrated at 13 mg/ml and were grown for 2 months at 20 °C. Crystals were fished, cryoprotected with 30% glycerol, and flash frozen in liquid nitrogen. Best crystal diffracted to 2.60 Å at ID14-EH1 beamline (Electrosynchrotron Research Facility, Grenoble). Data were processed using MOSFLM (26) and scaled using SCALA (27). Molecular replacement was performed using PHASER (28) included in the PHENIX package (29) and using the structure of NEP as the search model (Protein Data Bank ID code 1DMT). Automatic model building was performed using AUTOBUILD of the PHENIX suite, allowing the fitting of 90% of residues. The remaining residues were added manually.

Structure and sequence alignments relative to Fig. 1 were performed using 3DCoffee (30) and ClustalW (31), respectively, and edited with ESPript (32).

Structure was refined using COOT (33) and REFMAC (34) by applying TLS refinement to model data anisotropy. Zmp1 structure was analyzed and validated using MOLPROBITY (35). All molecular graphics images were produced using PyMOL (36) with the exception of Fig. 2*C*, prepared using the UCSF Chimera package (37).

*PDB Deposition* - The coordinates and the structure factors were deposited in the Protein Data Bank under ID code 3ZUK.

*Determination of Phosphoramidon Inhibition Constant* - Phosphoramidon inhibition constant $K_i$ was determined by following the inhibition of Zmp1 catalytic activity toward the generic fluorogenic substrate for matrix metalloproteinases (MMPs) MMP2/MMP7. This substrate consists of the amino acid sequence PLGL flanked by the fluorophore/quencher system methoxycoumarin/dinitrophenyl-alanine-arginine (full sequence: MCA-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-$NH_2$; MCA stands for (7-methoxycoumarin-4-yl)-acetyl; Dpa stands for *N*-3-(2,4-dinitrophenyl)-l-2,3-diaminopropionyl. MCA is the fluorophore, Dpa the quencher (MMP2/MMP7 product number 03-32-5032, Calbiochem). Several phosphoramidon concentrations, ranging from 15 to 250 nM, were incubated with 0.5 nM Zmp1 in 100 mM Tris-HCl, 0.1% borate, 150 mM NaCl, and 10 mM $CaCl_2$, pH 8.0, buffer at 37 °C. The fluorescence was detected in an Eclypse fluorometer (Varian) (excitation 320 nm, emission 395 nm). The inhibition constant $K_i$ was calculated upon determination of the various $K_m$ values from the substrate concentration dependence at several inhibitor concentrations and then plotting data of $K_m$ as a function of inhibitor concentration, as from the following equation.

$$K_m = {}^0K_m \cdot \left(1 + \frac{[I]}{K_I}\right) \qquad \text{(Eq. 1)}$$

## RESULTS

### Zmp1 Overall Structure and Active Site

Zmp1 was recombinantly expressed in *E. coli*, purified to homogeneity and co-crystallized with the inhibitor phosphoramidon, a metabolite of *Streptomyces tanashiensis* and a known broad spectrum inhibitor of Zn-dependent metalloproteases (38, 39). This molecule is able to inhibit in a competitive fashion (at least up to a concentration of 0.25 μM) the enzymatic activity of Zmp1 toward the synthetic substrate MMP2/MMP7 with a $K_i$ of 35 ± 5 nM (see supplemental Fig. 1). Orthorhombic crystals diffracted to 2.60 Å and contained two Zmp1 molecules in the asymmetric unit. The structure of Zmp1 was solved by molecular replacement using NEP as the search model. Automatic and manual model building led to a refined model with $R$ and $R_{free}$ factors converging to 17.51 and 24,96, respectively. Data

collection and refinement statistics are summarized in Table 1. Zmp1 shares 31% homology and 48% similarity with both human NEP and ECE-1 (Fig. 1) onto which it can be superimposed with a root mean square deviation of 1.87 Å and 1.73 Å for 593 and 594 Cα pairs, respectively. Zmp1 has two cysteines ($Cys^{285}$ and $Cys^{629}$), whereas NEP has 12 and ECE-1 14. $Cys^{428}$ of human ECE-1 ($Cys^{412}$ in rat ECE-1) is involved *in vivo* in the formation of an intermolecular disulfide bond responsible for ECE-1 dimerization (40). Zmp1, like NEP, is monomeric in solution (Ref. 24 and present work; data not shown), and its two cysteines are not involved in a disulfide bridge.

**TABLE 1**
**Data collection and refinement statistics of Zmp1**

| | |
|---|---|
| **Data collection statistics** | |
| Beamline | ESRF ID14-EH1 |
| Space group | $P2_12_12$ |
| Unit cell dimensions (Å), (°) | $a = 128.3$, $b = 198.8$, $c = 63.3$<br>$\alpha = \beta = \gamma = 90.00$ |
| Resolution (Å) | 30-2.60 (2.74-2.60) |
| Completeness (%) | 95.0 (75.6) |
| Multiplicity | 3.6 (2.3) |
| Observations | 175,191 (12,734) |
| Unique reflections | 48,222 (5,510) |
| $I/\sigma(I)$ | 12.2 (2.6) |
| $R_{merge}$ | 0.10 (0.37) |
| Molecules per asymmetric unit | 2 |
| Solvent content (%) | 54.9 |
| Wilson *B* factor ($A^2$) | 40.4 |
| **Refinement statistics** | |
| $R/R_{free}$ (%) | 17.51/24.96 |
| Root mean square deviation | |
| Bonds (Å)/Angles (°) | 0.015/1.52 |
| Ramachandran plot | |
| Residues in preferred regions (%) (no. of residues) | 95.8 (1,251) |
| Residues in allowed regions (%) (no. of residues) | 4.1 (53) |
| Residues in disallowed regions (%) (no. of residues) | 0.15 (2) |

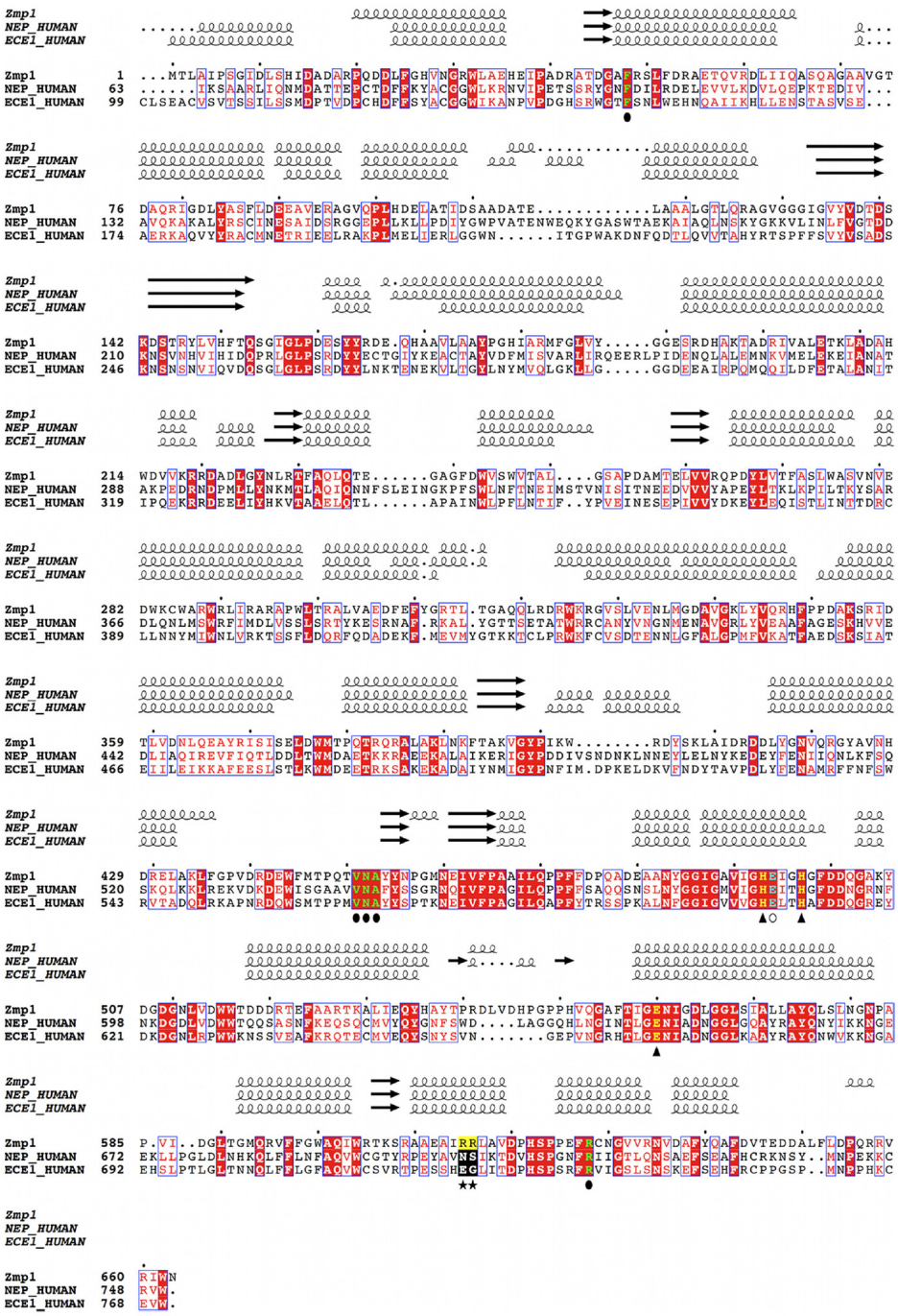



FIGURE 1. **Structure and sequence alignments of *M. tuberculosis* Zmp1 with human NEP and ECE-1.** *Spirals* and *arrows* indicate helices and β-strands, respectively. *Red boxes* with *white characters* indicate residue identity; *red characters* indicate residue similarity; *blue-framed characters* indicate similarities between groups of residues. Residues that bind $Zn^{2+}$ are shown in *yellow* and labeled with a *triangle*. Residues that interact with phosphoramidon are shown in *green* and labeled with a *black circle*. Catalytic $Glu^{494}$ is shown in *cyan* and labeled with a *white circle*. The two Arg residues of Zmp1 that make contact with TEA are colored *black* on a *yellow background* and are labeled with a *star*; the corresponding nonconserved NEP and ECE-1 residues are colored *white* on a *black background*.

The structure of Zmp1 presents an oval shape with dimensions of about 78 Å for the major axis and 60 Å for the minor one. Overall, the structure is composed by two mainly α-helical lobes interconnected by several loops distributed over the protein equatorial line (Fig. 2, *A* and *B*). The enzyme catalytic site is located between the two lobes and is accessible via two oppositely positioned small openings on the protein surface (Fig. 2*C*). The catalytic zinc ion is coordinated in a tetrahedral geometry by the conserved residues $His^{493}$ (2.2 Å) and $His^{497}$ (2.0 Å), belonging to the signature motifs $^{493}$**H**E*XX*$\mathbf{H}^{497}$, by $Glu^{560}$ (2.0 Å), part of the $^{560}$**E***XXX*$D^{564}$ signature fingerprint, and by the O1 oxygen of the *N*-phosphoryl moiety of the inhibitor (2.0 Å, Fig. 3*A*). The conserved residue $Glu^{494}$ (H$\mathbf{E}^{494}$*XX*H), unlike $Glu^{560}$, is not involved in metal coordination. Instead, its OE1 atom forms a specific interaction with the $O_2$ of the phosphoramidon *N*-phosphoryl group. This oxygen atom occupies the position of the water molecule responsible for the nucleophilic attack to the peptide substrate, as observed for other M13 family enzymes (23). Thus, the substitution of the catalytic water molecule with the phosphoramidon n-phosphoryl $O_2$ atom allows inhibition of Zmp1 catalysis.

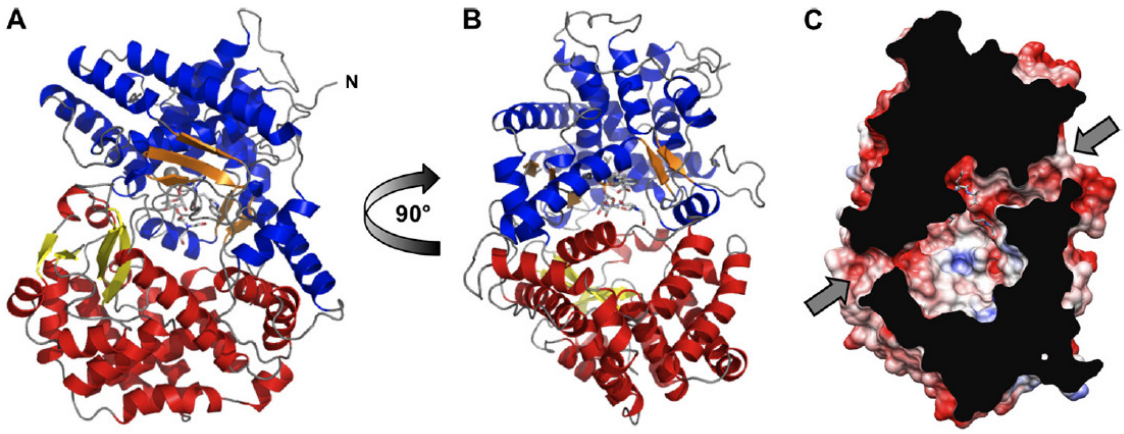



FIGURE 2. *A*, overall structure of Zmp1. *Colors* are coded according to secondary structure and lobe location. *Blue* and *red*, α-helices, upper and lower lobe, respectively; *orange* and *yellow*, β-sheets, upper and lower lobe, respectively; *dark gray*, loops. Phosphoramidon and TEA are shown in *sticks*. $Zn^{2+}$ is shown as a *gray sphere*. *B*, rotation of Zmp1 90 °C clockwise along the vertical axis. *C*, section of Zmp1 showing electrostatic surface potential. In *blue*, positive charges; in *red*, negative charges. *Arrows* indicate surface holes and cavities for product/substrate access to the catalytic site.

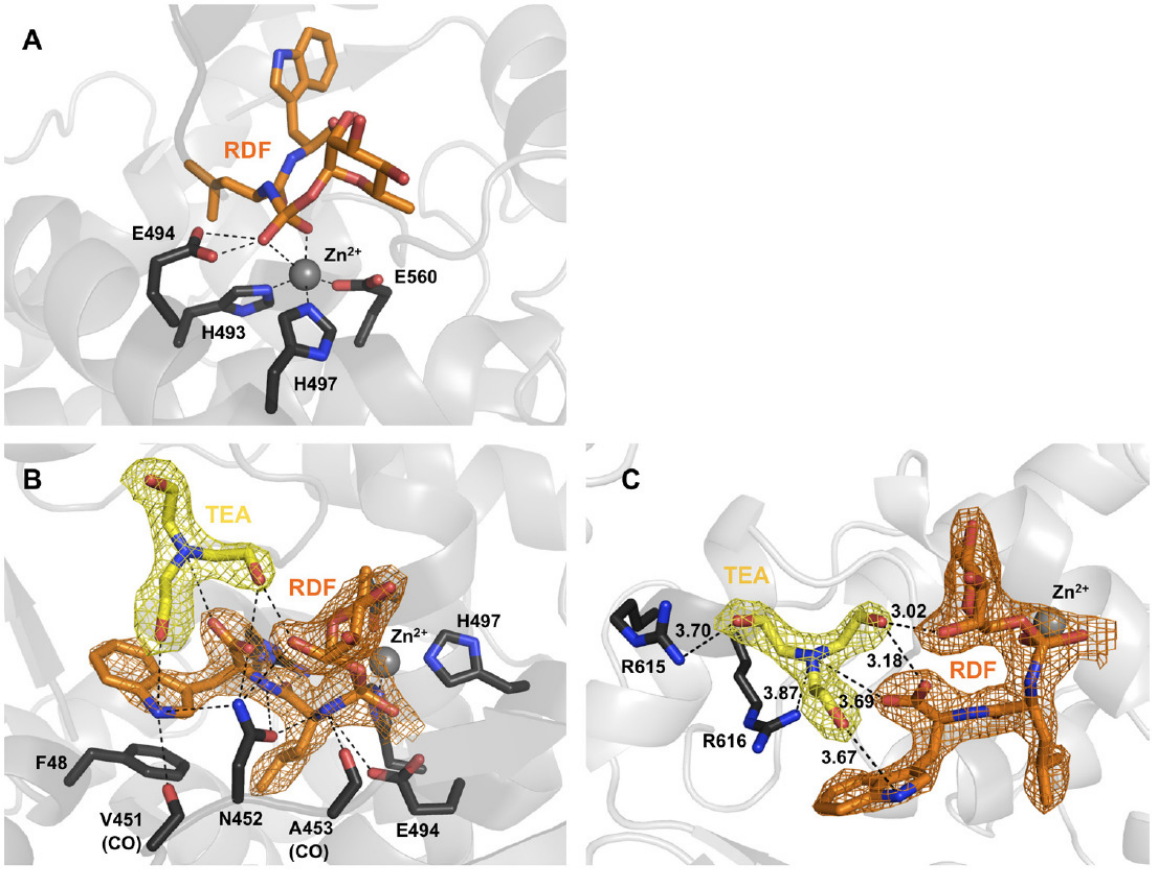



FIGURE 3. **Interactions of Zmp1 (chain A) with zinc (*A*) phosphoramidon (*B*), and TEA (*C*).** Zmp1 residues are rendered in *dark gray sticks*. Phosphoramidon and TEA and the corresponding electron densities are colored *orange* and *yellow*, respectively. Electron densities are contoured at 1.4 σ.

### Zmp1 Interactions with the Inhibitor Phosphoramidon

In the catalytic site, electron density compatible with the inhibitor molecule phosphoramidon was observed. Phosphoramidon is held in place by an intricate network of interactions with Zmp1 (Fig. 3). The indole moiety of phosphoramidon l-tryptophan makes π-stacking interaction with Phe$^{48}$, whereas the indole NH group forms a hydrogen bond with Val$^{451}$ main chain carbonyl group and Asn$^{452}$(ND2), both belonging to the conserved triad $^{451}$VNA$^{453}$ (Fig. 3*B*). OD1 and ND2 groups of Asn$^{452}$ are hydrogen-bonded with the inhibitor l-tryptophan and l-leucyl NH groups. The latter is also involved in hydrogen bonding with Ala$^{453}$ main chain carbonyl group and Glu$^{494}$ OE2. Additionally, the conserved Arg$^{628}$ forms hydrogen bonds with main chain l-aspartyl carbonyl group. The inhibitor rhamnose moiety does not make contacts with the protein and is exposed to the solvent.

**Zmp1 Interactions with N,N′,N''-Triethanolamine (TEA)**

Further electron density has been detected in close proximity to the phosphoramidon molecule and has been interpreted as TEA (Fig. 3, *B* and *C*). Because TEA was not added to any buffer used along the purification and crystallization procedures, we speculate it could likely be a contaminant of the phosphoramidon preparation. This molecule is observed only in one of the two Zmp1 subunits of the asymmetric unit (chain A) whereas in the other (chain B) is replaced by three water molecules matching the positions of the TEA hydroxyl groups of chain A. The TEA binding pocket is located within the catalytic site and the molecule forms stabilizing interactions both with the protein and the inhibitor phosphoramidon. All three hydroxyl groups of TEA make hydrogen bonds with both phosphoramidon and Zmp1 side chains: O3 with the $NH_2$ of the nonconserved $Arg^{615}$, O7 with NH and the carboxylic group of the l-tryptophan, and O10 with $O_2$ of the phosphoramidon rhamnose moiety. Additionally, the central amine group of TEA makes a hydrogen bond with the nonconserved $Arg^{616}$ (NH1).

**Zmp1 Recognition Subsites**

The S1 recognition subsite of Zmp1 accommodates the rhamnose moiety of the inhibitor (Fig. 4, *A* and *B*). Residues $Tyr^{454}$ and $Met^{446}$ do not interact with the phosphoramidon rhamnose group, leaving it mostly exposed to the solvent. These residues are conserved in ECE-1 ($Tyr^{568}$ and $Met^{560}$), whereas in NEP they are replaced by $Phe^{544}$ and $Ser^{536}$. The S1′ subsite of Zmp1 is constituted mainly by hydrophobic residues, as observed also in NEP and ECE-1 (Fig. 4, *C* and *D*). In Zmp1, residue $Ala^{489}$ ($Val^{603}$ in ECE-1) is replaced in NEP by a $Met^{579}$, making the Zmp1 S1′ subsite more hydrophobic compared with that of NEP. The indole moiety of the inhibitor makes π-stacking interaction with the conserved $Phe^{48}$ of Zmp1 ($Phe^{106}$ and $Phe^{149}$ in NEP and ECE-1, respectively). All of the other residues of the S1′ subsite ($Val^{490}$, $Ile^{468}$, $Phe^{473}$, $Ile^{603}$, and $Trp^{604}$) do not interact with the inhibitor and form a mostly hydrophobic pocket that presumably accommodates a substrate with a large, hydrophobic P1′ side chain.

The large S2′ subsite of Zmp1 is the least conserved among that of NEP and ECE-1 (Fig. 4, *E* and *F*). Only two residues of Zmp1 are conserved in NEP and ECE-1: $Phe^{48}$ and $Trp^{604}$ ($Phe^{106}$/$Phe^{149}$ and $Trp^{693}$/$Trp^{714}$ in NEP/ECE-1, respectively). In Zmp1,

the hydrophobic Phe$^{52}$ residue is replaced in NEP by the charged, flexible residue Arg$^{110}$ and in ECE-1 by the bulky residue Trp$^{153}$. Zmp1 residues Arg$^{45}$ and Asp$^{49}$ are substituted in NEP by Tyr$^{103}$ and Asp$^{107}$, respectively. The small, charged Zmp1 residues Thr$^{44}$ and Thr$^{606}$ are replaced in NEP by Arg$^{102}$ and Gly$^{695}$. Finally, Ser$^{608}$ is changed in Tyr$^{697}$ in NEP and in Arg$^{718}$ in ECE-1. All of these mutations not only alter the local charge and hydrophobicity, but also vary the accessible volume of the subsite. The differences among Zmp1, NEP, and ECE-1 S2′ subsites may be relevant for the specificity of their respective natural substrate(s).

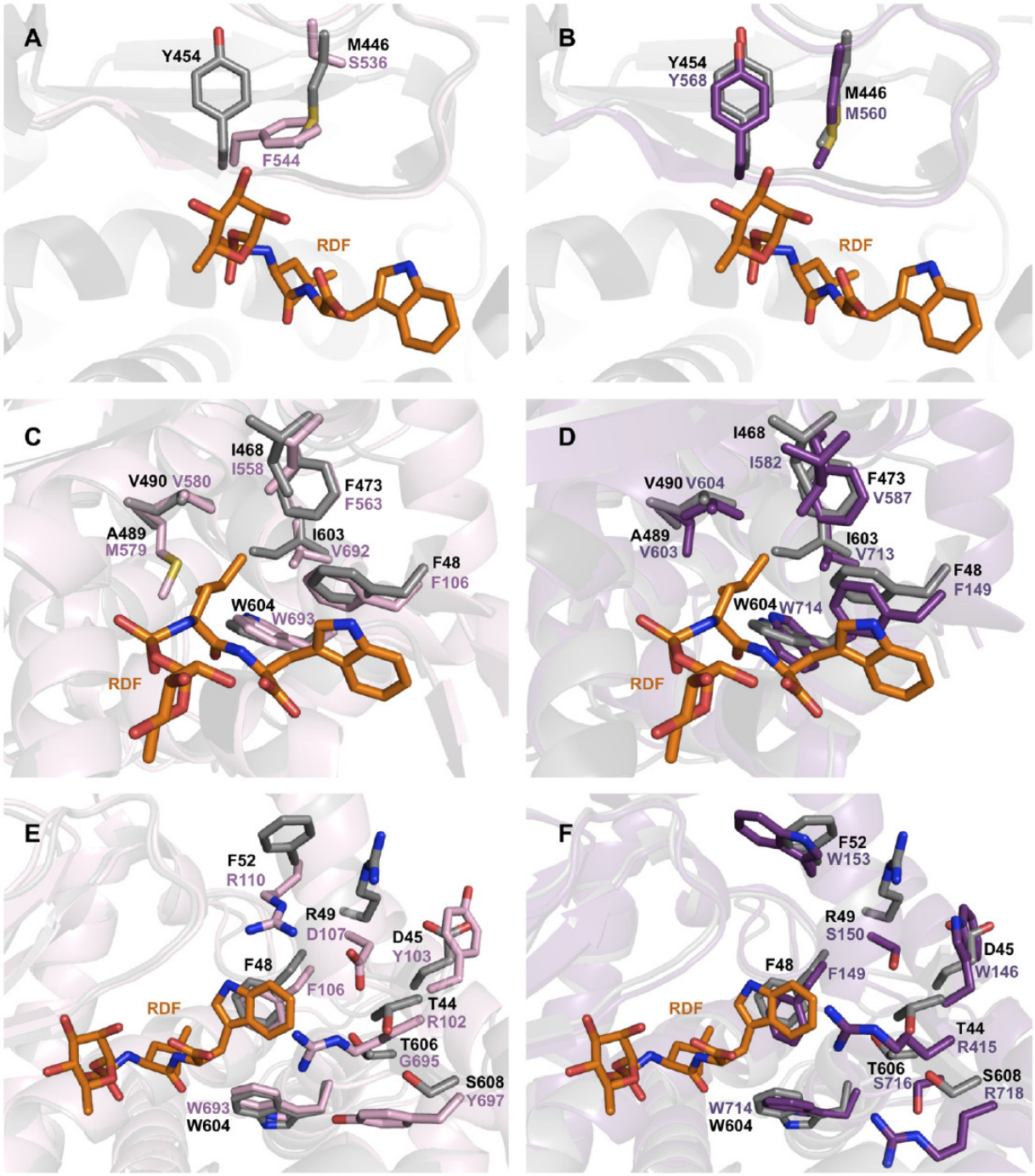



FIGURE 4. **Comparison of Zmp1 S1, S1′, and S2′ subsites with NEP and ECE-1.** *A*, Zmp1 (*gray*) and NEP (*pink*) S1 subsite. *B*, Zmp1 (*gray*) and ECE-1 (*violet*) S1 subsite. *C*, Zmp1 (*gray*) and NEP (*pink*) S1′ subsite. *D*, Zmp1 (*gray*) and ECE-1 (*violet*) S1′ subsite. *E*, Zmp1 (*gray*) and NEP (*pink*) S2′ subsite. *F*, Zmp1 (*gray*) and ECE-1 (*violet*) S2′ subsite. Phosphoramidon molecule is colored in *orange*.

## DISCUSSION

We report here the crystal structure of Zmp1 from *M. tuberculosis*, a potential pharmacological target against tuberculosis. It is to our knowledge the first crystal structure of a prokaryotic M13 protease. Zmp1 shows significant structural similarity with the human homologous proteins NEP and ECE-1 regarding the overall molecular architecture, the structural arrangement, and the mode of binding observed in complex with the inhibitor phosphoramidon. Residues forming the S1 subsite of Zmp1 are identical to those of ECE-1 but differ from those of NEP ($Tyr^{454}$ and $Met^{446}$ of Zmp1 are replaced by $Phe^{544}$ and $Ser^{536}$ in NEP). The local increased hydrophobicity of Zmp1 compared with the NEP S1 subsite could indicate a higher affinity for inhibitors carrying hydrophobic groups at the P1 site, as observed for ECE-1 (41, 42). However, the S1 site leaves the rhamnose moiety of the inhibitor exposed to the solvent and plays a minor role in substrate selectivity, as reported for NEP and ECE-1 (43–46). The S1′ subsite is highly conserved among Zmp1, NEP, and ECE-1. The residues present in this subsite are all hydrophobic with the exception, in NEP, of $Met^{579}$ ($Ala^{489}$/$Val^{603}$ in Zmp1/ECE-1, respectively). The slight difference in volume and hydrophobicity of the S1′ site resulting from this single mutation may explain the moderate preference, for ECE-1, of inhibitors carrying a bulky hydrophobic group at the P1′ site (1, 47, 48). A similar behavior could be envisaged also for Zmp1, as its S1′ subsite is entirely hydrophobic. Residues of the S2′ subsite are less conserved compared with those of the S1 and S1′ subsites. Understanding the mechanism of S2′ subsite specificity would be useful for the individuation of structural criteria that would serve as the rational basis for the design of specific Zmp1 inhibitors. The Zmp1 structure presented here bears in the catalytic pocket a molecule that we identified as TEA. In the crystal structure of NEP (24), a glycerol molecule occupies a position equivalent to that observed for TEA in Zmp1 and, like TEA, interacts both with NEP and phosphoramidon. Because TEA and glycerol are two structurally different molecules that bind to an equivalent pocket, we speculate that such a structural trait might be of relevance in the family of M13 metalloproteases for binding/recognition of nonsubstrate molecules possibly with regulatory role. It should also be noticed that such a secondary binding pocket shows

peculiar features in different members of the M13 family (Fig. 1). In particular, in Zmp1, TEA interacts with two Arg residues ($Arg^{615}$ and $Arg^{616}$) that are structurally nonconserved in human NEP and ECE-1-homologous proteins (Fig. 1). Therefore, the TEA binding pocket emerges as a promising docking site for the structure-based design of specific Zmp1 inhibitors. A matter of recent debate is the role played by Zmp1 in *M. tuberculosis* pathogenesis. A recent paper (17) shows that *zmp1*-deleted *M. tuberculosis* strains are hypervirulent, which contrasts with the work of Master *et al.* (12), as they propose that *zmp1* deletion leads to virulence attenuation. Although the issue awaits a definitive answer, the structure of Zmp1 reported here will help the developing of specific inhibitors that may prove useful both as a tool for further investigation of the biological functions of Zmp1 and as compounds of potential pharmaceutical interest.

**ACKNOWLEDGMENTS**

We thank the European Synchrotron Radiations Facility for providing synchrotron radiation facilities and Dr. Silvia Garavaglia (University of Piemonte Orientale) for help during data collection. This work was supported by the European Union FP7 program SysteMTb HEALTH-F4-2010-241587 and New TBvac 241745 projects and by Swiss National Science Foundation Grant 31003A_135705.

**REFERENCES**

1. De Lombaert, S., Blanchard, L., Stamford, L. B., Tan, J., Wallace, E. M., Satoh, Y., Fitt, J., Hoyer, D., Simonsbergen, D., Moliterni, J., Marcopoulos, N., Savage, P., Chou, M., Trapani, A. J., and Jeng, A. Y. (2000) *J. Med. Chem.* 43, 488–504

2. Koul, A., Arnoult, E., Lounis, N., Guillemont, J., and Andries, K. (2011) *Nature* 469, 483–490

3. Cooper, A. M. (2009) *Annu. Rev. Immunol.* 27, 393–422

4. Russell, D. G. (2007) *Nat. Rev. Microbiol.* 5, 39–47

5. Chan, J., and Flynn, J. (2004) *Clin. Immunol.* 110, 2–12

6. Armstrong, J. A., and Hart, P. D. (1971) *J. Exp. Med.* 134, 713–740

7. Kaufmann, S. H. (2001) *Nat. Rev. Immunol.* 1, 20–30

8. Russell, D. G. (2001) *Nat. Rev. Mol. Cell Biol.* 2, 569–577

9. Pieters, J. (2008) *Cell Host Microbe* 3, 399–407

10. Ramachandra, L., Smialek, J. L., Shank, S. S., Convery, M., Boom, W. H., and Harding, C. V. (2005) *Infect. Immun.* 73, 1097–1105

11. Torres, M., Ramachandra, L., Rojas, R. E., Bobadilla, K., Thomas, J., Canaday, D. H., Harding, C. V., and Boom, W. H. (2006) *Infect. Immun.* 74, 1621–1630

12. Master, S. S., Rampini, S. K., Davis, A. S., Keller, C., Ehlers, S., Springer, B., Timmins, G. S., Sander, P., and Deretic, V. (2008) *Cell Host Microbe* 3, 224–232

13. Sutterwala, F. S., Ogura, Y., and Flavell, R. A. (2007) *J. Leukocyte Biol.* 82, 259–264

14. Schroder, K., and Tschopp, J. (2010) *Cell* 140, 821–832

15. Guarda, G., and So, A. (2010) *Immunology* 130, 329–336

16. Dinarello, C. A. (2011) *Blood* 117, 3720–3732

17. Muttucumaru, D. G., Smith, D. A., McMinn, E. J., Reese, V., Coler, R. N., and Parish, T. (2011) *Tuberculosis* 91, 111–116

18. Bianchetti, L., Oudet, C., and Poch, O. (2002) *Proteins* 47, 481–488

19. Ikeda, K., Emoto, N., Raharjo, S. B., Nurhantari, Y., Saiki, K., Yokoyama, M., and Matsuo, M. (1999) *J. Biol. Chem.* 274, 32469–32477

20. Xu, D., Emoto, N., Giaid, A., Slaughter, C., Kaw, S., deWit, D., and Yanagisawa, M. (1994) *Cell* 78, 473–485

21. Lee, S., Zambas, E. D., Marsh, W. L., and Redman, C. M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 6353–6357

22. HYP-consortium (1995) *Nat. Genet.* 11, 130–136

23. Tallant, C., Marrero, A., and Gomis-Ru¨th, F. X. (2010) *Biochim. Biophys. Acta* 1803, 20–28

24. Oefner, C., D'Arcy, A., Hennig, M., Winkler, F. K., and Dale, G. E. (2000) *J. Mol. Biol.* 296, 341–349

25. Schulz, H., Dale, G. E., Karimi-Nejad, Y., and Oefner, C. (2009) *J. Mol. Biol.* 385, 178–187

26. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) *Acta Crystallogr. D Biol. Crystallogr.* 67, 271–281

27. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D Biol. Crystallogr.* 50, 760–763

28. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J. Appl. Crystallogr.* 40, 658–674

29. Adams, P. D., Afonine, P. V., Bunko´czi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., Mc-Coy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221

30. O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D. G., and Notredame, C. (2004) *J. Mol. Biol.* 340, 385–395

31. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., and Thompson, J. D. (2003) *Nucleic Acids Res.* 31, 3497–3500

32. Gouet, P., Courcelle, E., Stuart, D. I., and Me´toz, F. (1999) *Bioinformatics* 15, 305–308

33. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501

34. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) *Acta Crystallogr. D Biol. Crystallogr.* 60, 2184–2195

35. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) *Acta Crystallogr. D Biol. Crystallogr*. 66, 12–21

36. DeLano, W. L. (2010) *The PyMOL Molecular Graphics System*, DeLano Scientific LLC, San Carlos, CA

37. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) *J. Comput. Chem.* 25, 1605–1612

38. Suda, H., Aoyagi, T., Takeuchi, T., and Umezawa, H. (1973) *J. Antibiot.* 26, 621–623

39. Lorand, L. (1976) *Methods Enzymol.* 45, 31–37

40. Shimada, K., Takahashi, M., Turner, A. J., and Tanzawa, K. (1996) *Biochem. J.* 315, 863–867

41. Kukkola, P. J., Savage, P., Sakane, Y., Berry, J. C., Bilci, N. A., Ghai, R. D., and Jeng, A. Y. (1995) *J. Cardiovasc. Pharmacol.* 26, S65–68

42. Shetty, S. S., Savage, P., DelGrande, D., De Lombaert, S., and Jeng, A. Y. (1998) *J. Cardiovasc. Pharmacol.* 31, S68–70

43. Johnson, G. D., Swenson, H. R., Ramage, R., and Ahn, K. (2002) *Arch. Biochem. Biophys*. 398, 240–248

44. Gaucher, J. F., Selkti, M., Tiraboschi, G., Prange´, T., Roques, B. P., Tomas, A., and Fournie´-Zaluski, M. C. (1999) *Biochemistry* 38, 12569 –12576

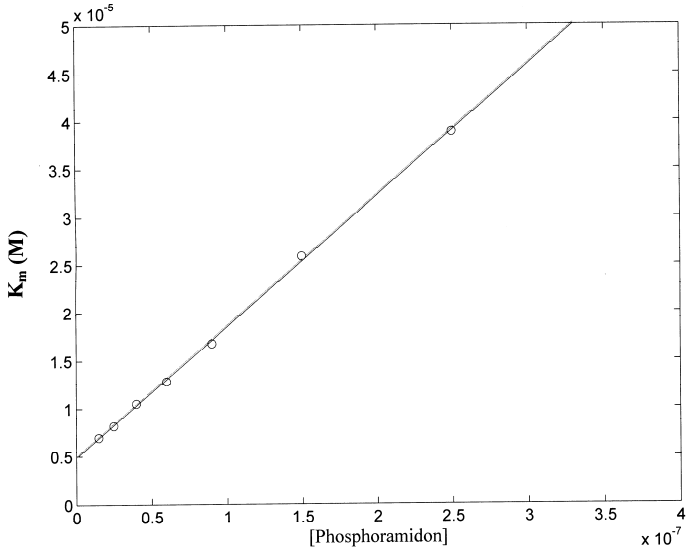
45. Ksander, G. M., de Jesus, R., Yuan, A., Ghai, R. D., McMartin, C., and Bohacek, R. (1997) *J. Med. Chem.* 40, 506–514

46. Ksander, G. M., de Jesus, R., Yuan, A., Ghai, R. D., Trapani, A., McMartin, C., and Bohacek, R. (1997) *J. Med. Chem.* 40, 495–505

47. Oefner, C., Roques, B. P., Fournie-Zaluski, M. C., and Dale, G. E. (2004) *Acta Crystallogr. D Biol. Crystallogr.* 60, 392–396

48. De Lombaert, S., Ghai, R. D., Jeng, A. Y., Trapani, A. J., and Webb, R. L. (1994) *Biochem. Biophys. Res. Commun.* 204, 407–412

## SUPPLEMENTAL DATA




**Supplementary Figure 1.** Phosphoramidon concentration dependence of MCA substrate affinity constant $K_m$ for Zmp1 at pH 8.0 and 37°C. Values of $K_m$ were obtained by Lineweaver-Burk plot of fluorogenic MCA peptide cleavage in the presence of 0.5 nM Zmp-1 in the absence and presence of different Phosphoramidon concentrations. The continuous line is the non-linear least-squares fitting of data according to Eq. (1), from which a value of ${}^0K_m = 4.8(\pm 0.7) \times 10^{-6}$ M and a value of $K_I = 3.5(\pm 0.5) \times 10^{-8}$ M are obtained.

# 3. Application of the proteomic approach 2D-PAGE for Zmp1 substrate identification

## INTRODUCTION

Proteomics is the large-scale, systematic study of proteins, particularly their functions, modifications, interactions and localization. Usually, the tools of proteomics are applied to investigate modifications induced during the course of cell differentiation, activation or transformation. The proteolysis, a key regulatory post-translational modification in diverse cellular processes, is a privileged model for this type of study. Indeed, proteomic approaches permit to track proteolytic events and identify protease substrates. Considering the functional relevance of proteases for all living processes, and the fact that many infectious microorganisms, viruses and parasites use proteases as virulence factors, there has been an increasing interest in developing protease research. In the past years, a new field emerged with the name 'degradomics', described as all genomic and proteomic approaches for the identification and characterization of proteases present in an organism, including the substrates that are targeted by these proteases and their endogenous inhibitors. The complete natural substrate repertoire of a protease in a cell, tissue or organism is then called 'degradome' (Lopez-Otin and Overall, 2002).

A conventional gel-based proteomic method is the two dimension gel electrophoresis (2D-PAGE), which can be used for comparison and relative quantification of various proteome states (Issaq and Veenstra, 2008). This technique separates proteins according to two independent proprieties in two discrete steps. The first dimension step, isoelectric focusing, separates proteins according to their isoelectric points; the second dimension step, sodium dodecyl sulfate-polyacrylamide gel eletrophoresis (SDS-PAGE), separates proteins according to their molecular weight. As result, each spot on the gel corresponds to a single protein species in the sample. The coupling of this method with mass spectrometry allows the identification of proteins. Despite alternative technologies that have emerged, 2D-gel is currently the only technique that can routinely separate up to thousands of proteins in complex protein mixtures. Furthermore, it delivers a map of intact proteins that reflects changes in protein expression level, post-translation modifications, and isoforms. When first used, this

approach enabled the discovery of protease substrates (Hwang *et al.*, 2004). Using this method indeed, the protease is able to interact with its substrates in their native, folded conformations, mimicking the interaction of enzyme and substrate *in vivo*. Natural substrates are then cleaved by the protease, while remaining intact in a control sample not treated with protease. To determine which of the many proteins in the complex mixture were cleaved by the protease, the 2D protein profiles of control and protease-treated samples are compared. Protein spots that are in greater abundance in the control sample than the protease-treated sample are candidate substrates that underwent cleavage. On the other hand, spots that are in greater abundance in the protease-treated sample are potential cleavage products generated by the protease (Fig. 1).

2D-gel methods have been applied to caspase substrate identification, revealing 41, 13, and 15 putative substrates of caspase-1, caspase-3, and caspase-7, respectively (Agard and Wells, 2009). Similarly, the degradome of the bacterial protease Isp-1 (Lee *et al.*, 2004) was examined and substrates of the matrix metalloprotease-14 have been identified in human plasma (Hwang *et al.*, 2004). Moreover, its application for proteome mapping contributed to identify novel proteins involved in specific cellular pathways or compartments. The Proteome 2D-PAGE Database is a database for storing 2D-gel maps of a broad range of organisms, from human to microorganisms (http://web.mpiib-berlin.mpg.de/cgi-bin/pdbs/2d-page/extern/index.cgi). 2D-PAGE maps of *M. tuberculosis* H37Rv culture filtrate and lysate proteins have been established, as well as the maps of the cellular proteins of *M. bovis* BCG substrain Copenhagen and Chicago. 231 and 138 different *M. tuberculosis* ORFs have been identified in 2D-PAGE maps of the cellular proteins and culture supernatant proteins, respectively.

Due to its large and successful application in protease substrate identification, the 2D-PAGE coupled to mass spectrometry was employed to identify Zmp1 substrates. The main goal of this study was to identify putative endogenous substrates of the zinc metallopeptidase Zmp1. Although Zmp1 was found in *M. tuberculosis* supernatant (Master *et al.*, 2008), in this study the protein was detected uniquely in the cell pellet. Hence, the 2D-PAGE maps of the proteome of BCG WT and the mutant BCG Δzmp1 were compared. Moreover, a comparative analysis of the

proteome of BCG Δzmp1 incubated in presence or absence of purified active Zmp1 was performed. In both approaches, the protein profiles were investigated and the differences in the spot intensity were analysed.

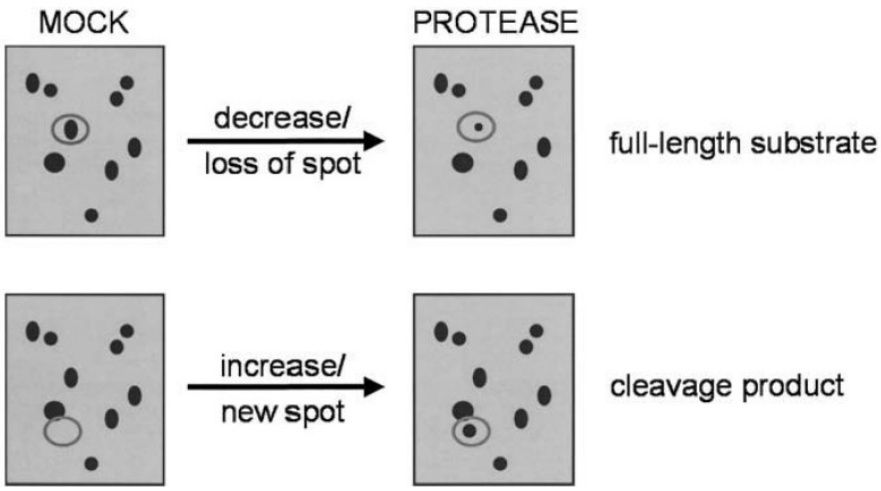



Fig.1 Scheme of 2D-PAGE applied for protease substrate identification. The circled spot (top) shows a decreased intensity in the protein profile of the treated sample, indicating cleavage event on that protein. Conversely, substrates may be identified by detection of their cleavage products, which appear as new spot in the treated sample (bottom). Both full-length proteins and cleavage products can be picked and identified by MS (Bredemeyer *et al.*, 2005).

## EXPERIMENTAL PROCEDURES

### Bacterial strains and protein extraction

*M. bovis* BCG Pasteur WT and *M. bovis* BCG Pasteur Δ*zmp1* (Master *et al.*, 2008) were cultured in 7H9 medium or Sauton's medium at 37° C, cells were harvested by centrifugation in late exponential phase (3 weeks). For preparation of the lysate, bacteria were washed twice in PBS, pH 7.4. The bacterial pellet was resuspended in PBS, and sonicated for 15 min at 4°C in the sonication bath (Transsonic 460, Elma). The total protein extract resulted from the supernatant recovered after centrifugation of the homogenates at 16000 g for 20 min at 4° C. The protein concentration was determined by BCA protein assay (bicinchoninic acid).

**Western blot analysis**

50 µg of total protein extracts and 1-2 ml of precipitated culture supernatant were loaded on SDS-PAGE (12.5 % Tris-HCl gels, *BioRad*) and immunoblotted on a PVDF membrane (*BioRad*) by a semi dry apparatus (*BioRad*). The membranes were incubated in antibody solutions against Zmp1 (1:2500), KatG (1:5000) and Antigen85 (1:2500). A second antibody coupled with HRP (horseradish peroxidase, 1:20000, *Invitrogen*) was used to visualise the first antibody. Zmp1 antibody was a gift from V. Deretic.

**Proteolytic digestion with the zinc metallopeptidase Zmp1**

The zinc metallopeptidase Zmp1 was purified as described in Petrera *et al*., 2012. The peptidase was stored at – 80° C in aliquots of 1 mg/ml. A proteolytic digestion was set up in a final volume of 1 ml. 600 µg of protein extract were incubated with 30 µg of Zmp1 in the reaction buffer (100 mM TRIS pH 7.3, 100 mM NaCl, 0.05% BriJ) for 1h at 37° C. A parallel reaction with inactive Zmp1 (heat-inactivated at 80° C for 20 min) occurred. Zmp1 activity before and after proteolytic reaction was assessed with a fluorogenic substrate, as described in Petrera *et al*., 2012. After reaction, protein was precipitated with TCA-acetone. 13% ice cold acetone was added and the mixture was incubated for 2 h at 4° C. Precipitated protein was collected by centrifugation (16000 g, 20 min, 4° C), washed with one volume of 100% ice cold acetone, and dried by vacuum concentrator (Concentrator 5306, Eppendorf) for 5 min at RT.

**Two-dimensional gel electrophoresis Isoelectric Focusing (IEF)**

For isolectric focusing (IEF), rehydration buffer (7M Urea, 2M Thiourea, 2% CHAPS, 50 mM DTT, 2% Pharmalyte, 0.123% TBP) was added to the protein pellet. The resuspended protein mixture was split for running three parallel gels (200 µg/gel). 24 cm immobilized pH gradient (IPG) strips (Bio-Rad) of pH range 4-7 were rehydrated overnight at 20° C with 450 µl of sample containing protein (200 µg). IEF was performed in a IEF cell (Bio-Rad) using the following six-step program: a) 0-150 V, 3 h; b) 300 V constant for 3 h; c) 1000 V, 5 h; d) 8000 V, 3 h; e) 8000 V constant for 9 h; f) 500 V constant for minimum 1 h. The current limit was set at 50 µA/strip.

**SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)**

After IEF, each strip was equilibrated for 15 min in a solution containing 0.05 M Tris-HCl, pH 8.8, 6 M urea, 30 % glycerol, 2 % SDS, 130 mM DTT followed by a second

bath with the same solution but containing 135 mM iodoacetamide in place of DTT. Each IPG strip was placed on top of a vertical SDS-polyacrylamide gel (12.5% polyacrylamide concentration) and sealed in place with 0.5 % low melting agarose dissolved in running buffer (25 mM Tris-HCl, pH 8.3, 192 mM glycine, 0.1 % SDS) containing 0.002 % bromophenol blue. Molecular mass markers (Bio-Rad) were loaded in a separate well by the side of the strip. SDS-PAGE was performed in Ettan Dalt II system (Amersham Biosciences) with three-step program: a) 2 W/gel for 45 min; b) 17 W/gel for 4 h; c) 1 W/gel as a hold. After electrophoresis gels were stained with SYPRO Ruby and photographed.

**In-gel digestion and MS analysis**

The spot intensity and the intensity ratio WT/mutant were calculated with the programs ProteomWeaver and Progenesis SameSpots. Stained spots were excised from the gel, washed with 100 mM ammonium bicarbonate, dehydrated with acetonitrile (80% solution) for 10 min and dried in a vacuum concentrator for 5 min. The proteins were then digested with trypsin (Promega) for 3-6 h at 37° C. The tryptic peptides were prepared for MALDI-TOF analysis by concentration and desalting using ZipTip $_{C18}$ (Millipore). They were eluted with saturated α-cyano-4-hydroxy cinnamic acid solution in 50 % acetonitrile, 0.1 % TFA, and 0.7 µl of the sample were applied to the steel target plate and analysed by MALDI-TOF/MS and MALDI-MS/MS. The obtained mass spectra were searched against the *M.tuberculosis* complex database using MASCOT (http://www.matrixscience.com).

## RESULTS

### Study of localization of Zmp1

The question about the secretion of Zmp1 was previously answered by Master and colleagues (Master *et al.*, 2008). The protein was mainly found intracellularly and in minor amount in the supernatant of *M. tuberculosis* (Fig.2A). The abundant intracellular localization of Zmp1 in BCG WT was confirmed in this study. As shown in Figure 2B, Zmp1 was only detected in cell pellet of BCG WT, but not in the supernatant. Therefore, the bacterial proteome could contain Zmp1 substrates.

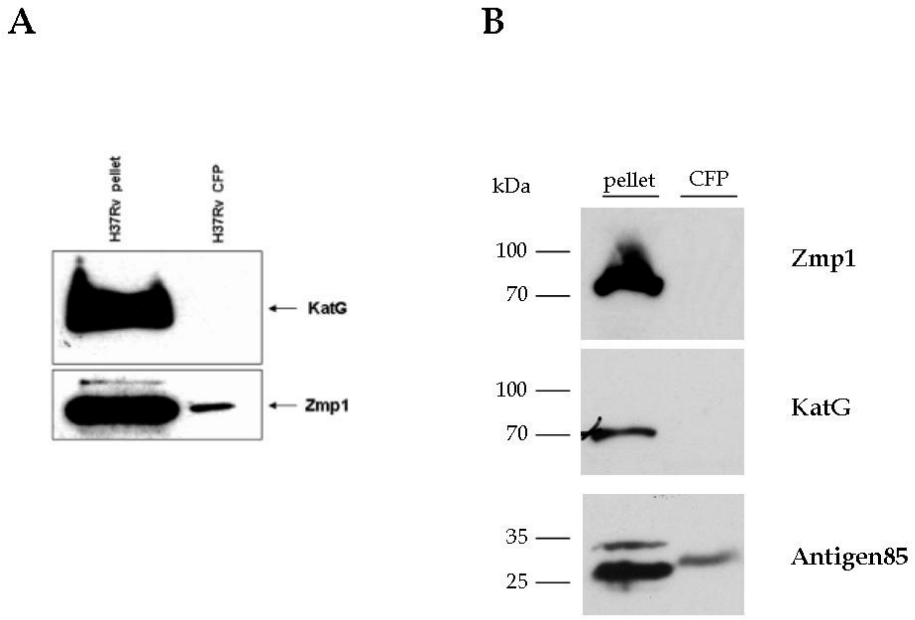



Fig. 2 A. Immunoblot analysis of Zmp1 in *M. tuberculosis* H37Rv in culture supernatants and cell pellet (Master *et al.*, 2008). Culture filtrate protein (CFP) and cell pellet were subjected to Western blotting and probed against antibodies specific for Zmp1 and KatG. The mycobacterial cytoplasmic enzyme KatG was probed for assessing nonspecific lysis in the CFP fraction. B. Immunoblot analysis of Zmp1 in *M. bovis* BCG WT in culture filtrate and cell pellet. The secreted mycobacterial Antigen85 was probed for assessing the preparation of supernatant proteins.

**Analysis and comparison of 2D-gel profiles of *M. bovis* BCG WT and *M. bovis* BCG Δzmp1**

A first approach dealt with the comparison of the proteome of BCG WT and BCG Δzmp1, searching for modification in the mycobacterial protein pattern due to the presence of the zinc metallopeptidase Zmp1. 600 µg of total protein extract were precipitated with TCA/acetone. The protein pellet was suspended in appropriate volume of 2D rehydration buffer and then split for running three parallel gels. Finally it was subjected to first-dimensional isoelectric focusing and second-dimension SDS-PAGE. The pH range 4-7 was selected as the majority of proteins have a pI within that range, permitting therefore to zoom into this region. The use of the immobilized pH gradient (IPG) technology offered significant advantage in terms of resolution and reproducibility. The original method for first-dimension IEF was based on carrier-ampholyte-generated pH gradient in cylindrical polyacrylamide

gels cast in glass rods or tubes. As a result of limitations and problems associated with carrier ampholyte pH gradients, an immobilized pH gradient was developed, by covalently incorporating a gradient of acidic and basic buffering groups (immobilines) into a polyacrylamide gel at the time it is cast. For the second dimension, 12.5 % polyacrylamide concentrated SDS-gel was chosen to detect proteins in the broad range 15-100 kDa. The 2D-gels of wild-type and mutant show a very similar protein pattern (Fig. 3). The intensity of the spots was quantified and 32 spots have been picked as recording an intensity ratio WT/mutant above 1.5 or below 0.6. 21 out of them were identified by MS analysis. The identified spots are listed in Table 1. The functional analysis of the hits revealed that the majority of the proteins are metabolic enzymes, involved in diverse cellular processes. The overrepresentation of this class of enzymes can be explained considering the two different genetic backgrounds here compared. The lack of the zinc metallopeptidase Zmp1 might influence the cellular metabolism and indirectly perturb some biological pathways. Therefore in two different genetic backgrounds some enzymes can be differently expressed. Although it is clearly interesting to study the modification of the protein expression pattern upon depletion of Zmp1, the subject of this study was focused on the identification of its substrates. The spot recording the highest ratio WT/mutant was identified as the conserved hypothetical protein Hrp1, hypoxic response protein 1, encoded by the *M. tuberculosis* gene Rv2626c (MW=15.5 kDa). The spot appears 3.7 fold more intense in the WT sample, i.e. in presence of Zmp1, suggesting that it is a cleavage product of the protein Hrp1. As the molecular mass of the spot corresponds approximately to the full size protein, it could be envisaged a partial C- or N-terminal digestion of Hrp1 by Zmp1. The second highest hit is the putative uncharacterized protein Rv1680. No information is available about this protein, although the large size (274 amino acids) decreases the chances of being cleaved by Zmp1, which shows a clear preference for small proteins (Ferraris *et al.*, 2011).

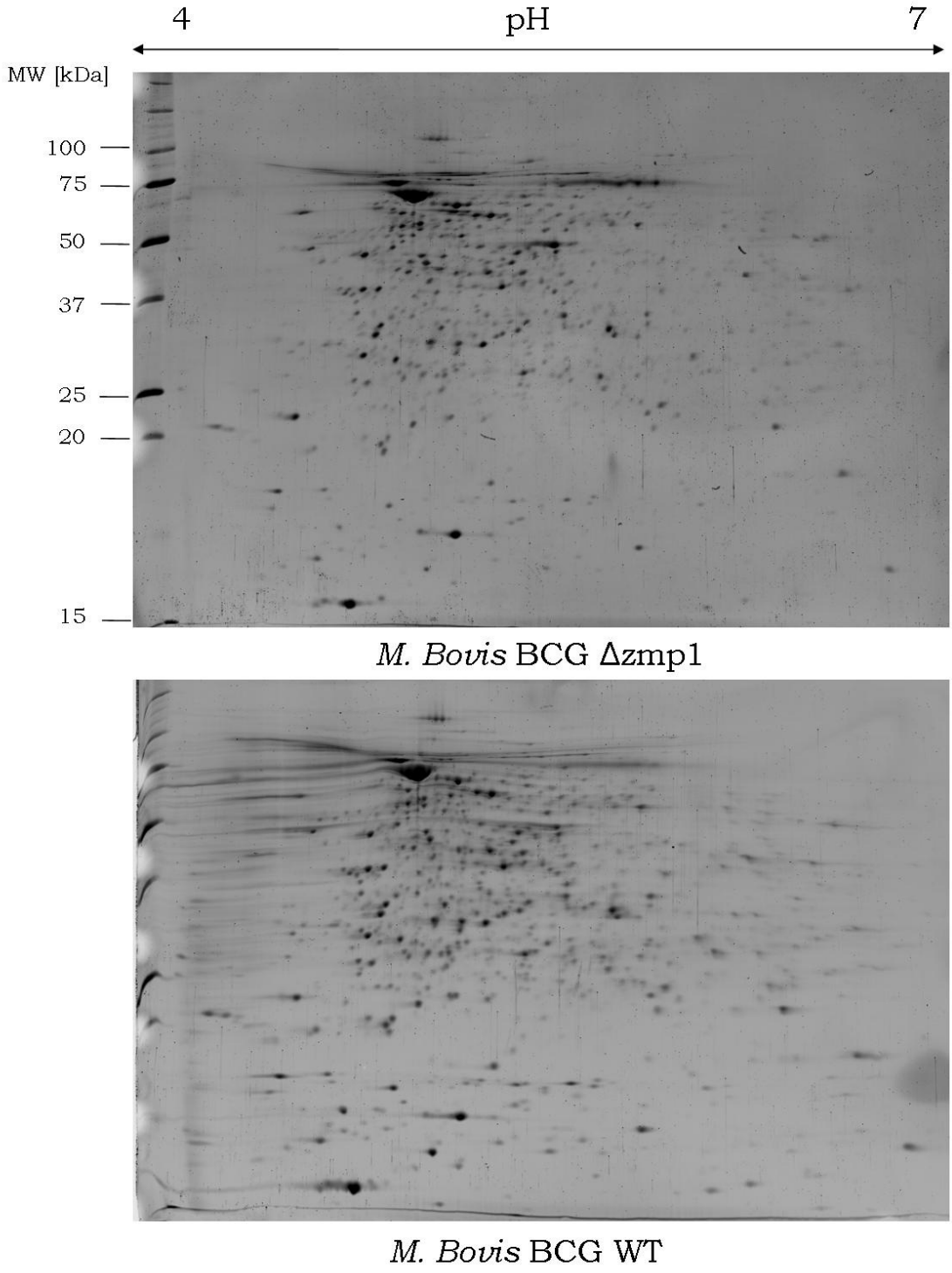



Fig.3 2D-PAGE analysis of BCG WT and BCG Δzmp1. Mycobacterial lysates were prepared from cultured strains. The proteins were separated on 2D gel, and then SYPRO Ruby-stained for visualization of protein spots. (top) 2D-map of BCG Δzmp1; (bottom) 2D-map of BCG WT. Here represented one of the three gels run for each strain.

| | Proteins overrepresented in WT | Fold WT/mutant |
|---|---|---|
| 1 | Hypoxic response protein 1 Rv2626c | 3.7 |
| 2 | Putative uncharacterized protein Rv1680 | 2.6 |
| 3 | Enoyl-[acyl-carrier-protein] reductase [NADH] inhA | 2.5 |
| 4 | Uncharacterized protein Mb1315 (Carbonic_anhydrase) | 2.3 |
| 5 | Putative uncharacterized protein Rv0801 (Glyoxalase Family) | 2.2 |
| 6 | Bacterioferritin A Mb1907 | 2.1 |
| 7 | Beta 1,3-Glucanase Rv0315 | 2 |
| 8 | Divalent cation-transport integral membrane protein mntH | 2 |
| 9 | Enolase Mb1051 | 2 |
| 10 | Riboflavin synthase alpha chain ribE | 1.9 |
| 11 | ATP-dependent Clp protease proteolytic subunit 1 | 1.8 |
| 12 | PhosphatidylEthanolamine-Binding Protein Mb2164c | 1.7 |
| | **Proteins underrepresented in WT** | |
| 13 | Isocitrate lyase Mb0476 | 0.28 |
| 14 | Fatty acid desaturase MT0846 | 0.32 |
| 15 | Possible dioxygenase Mb3186c | 0.32 |
| 16 | DNA-binding response regulator mtrA | 0.43 |
| 17 | Probable glutamine synthetase 2 Mb2246c | 0.43 |
| 18 | Elongation factor Tu | 0.5 |
| 19 | Succinyl-CoA synthetase beta chain | 0.52 |
| 20 | Uncharacterized protein (stress protein UspA) | 0.52 |
| 21 | Glutamine synthetase 1 glnA1 | 0.6 |

Table 1. List of identified proteins which recorded a significant intensity ratio between WT and mutant. The spot intensity ratio was calculated by the program ProteomWeaver. The values result from the analysis of three parallel gels.

### Analysis and comparison of 2D-gel profiles of *M. bovis* BCG Δzmp1 incubated with Zmp1

The disadvantage of dealing with two different backgrounds, inevitably introducing variations in protein expression pattern, was overcome by comparing the same bacterial proteome treated or untreated with purified active Zmp1. The mycobacterial proteins were extracted from BCG Δzmp1, in order to arrange a complete set of proteins not affected by Zmp1 activity. 600 µg of protein extract were incubated with 30 µg of active or heat-inactivated Zmp1 in the reaction buffer (see Experimental procedures) for 1 h at 37° C. The protein mixture was then precipitated, the pellet was resuspended in rehydration buffer and the volume was split for running three parallel gels. Comparison of 2D gel patterns of mycobacterial proteins

treated with active or inactivated Zmp1 demonstrated that Zmp1 treatment did not cause nonspecific, overall degradation of bacterial intracellular proteins. Instead, specifically six spots showed a decreased or increased intensity upon Zmp1 treatment (Fig.4). Three spots out of six were identified by MS analysis. Spot 1 is the riboflavin synthase alpha chain (intensity ratio WT/mutant = 2.1). Spot 4 is the bacterioferritin A (intensity ratio WT/mutant = 1.9). Both proteins were identified in the previous experiment (hits num. 10 and 6, respectively). The spot 6 is the protein Rv2626c (intensity ratio WT/mutant = 3). The spot appears also in this experiment more pronounced in presence of Zmp1 (Fig. 5A and 5B). To confirm Rv2626c as a substrate of Zmp1, the experiment was repeated. As shown in the Fig.5C, the result was reproduced. Two out of three spots have still decreased intensity in the control gels.

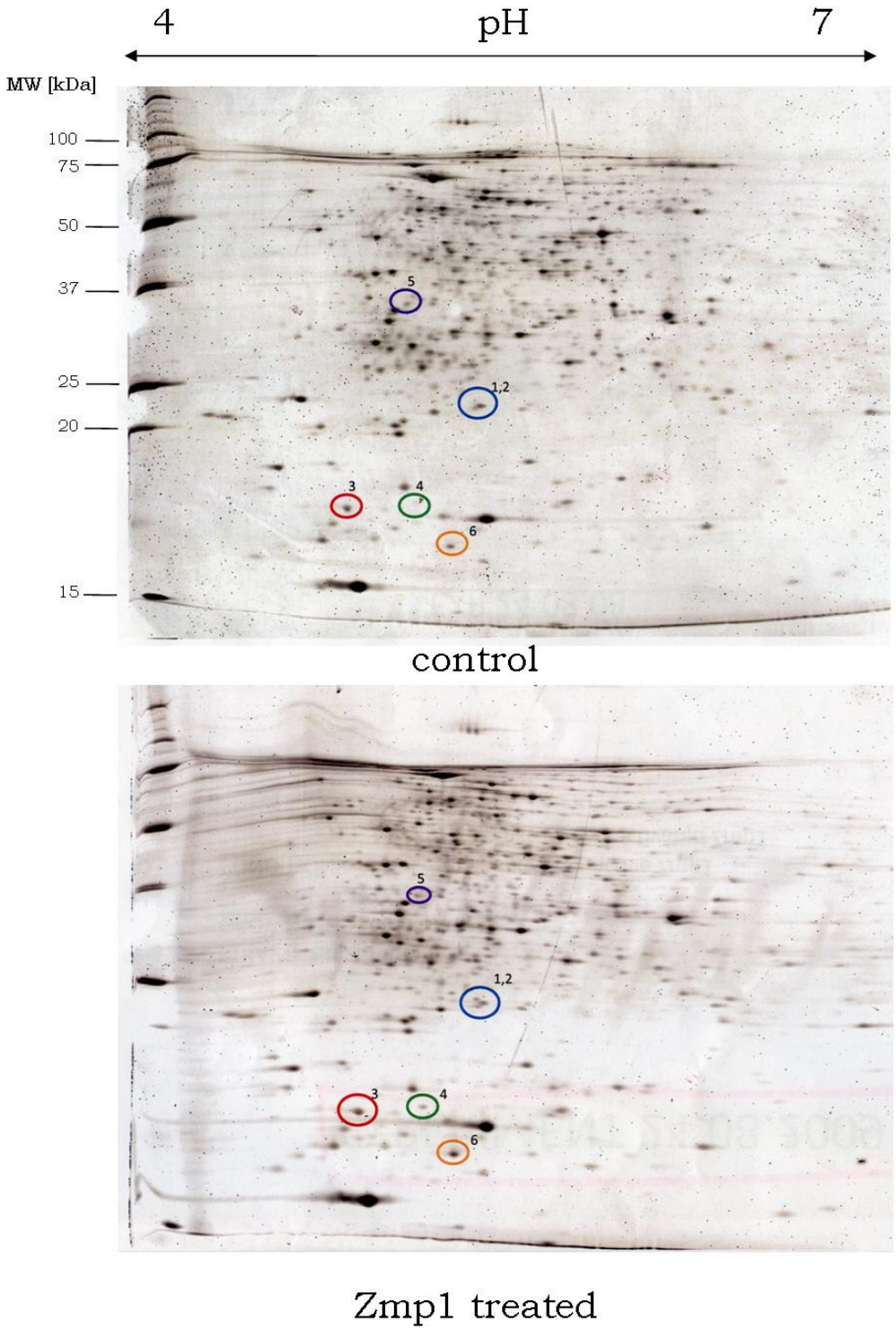



Fig.4 2D-PAGE analysis of *M. bovis* BCG Δzmp1 proteome incubated with Zmp1 for 1 h at 37° C. The proteome derives from the mutant strain *M. bovis* BCG Δzmp1, which after treatment with active and heat-inactivated Zmp1 was loaded in 2D-gel. The gel was stained with SYPRO Ruby for visualization of protein spots. (top) 2D-map of *M. bovis* BCG Δzmp1 untreated proteome; (bottom) 2D-map of *M. bovis* BCG Δzmp1 proteome treated with Zmp1. Here represented one of the three gels run for each condition.

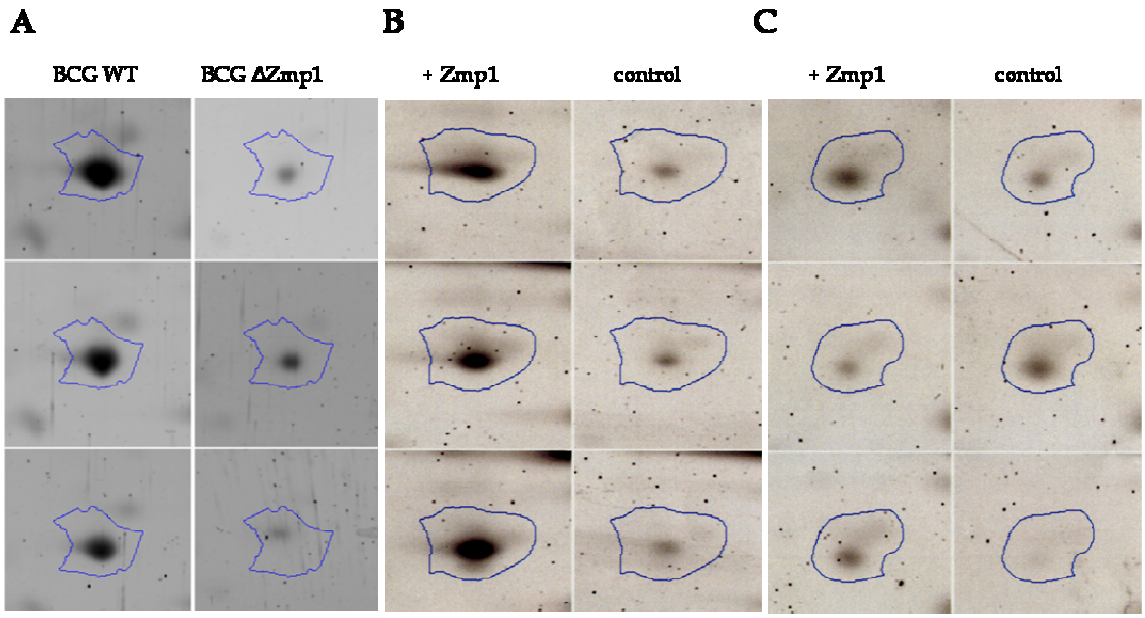



Fig.5 Spot of the protein Rv2626c. (A) The spot in the comparative analysis of the wild-type and mutant proteomes. (B,C) The spot in the 2D-gel of the mutant proteome incubated with Zmp1 in the first (B) and in the second trial (C).

## DISCUSSION

The search of Zmp1 substrates was approached by the application of proteomic tools. The method considered the most suitable for the aim of the study was the 2D-PAGE, as it turned out to be very fruitful in the investigation of several eukaryotic and prokaryotic protease substrates. This study started questioning which proteome is targeted by Zmp1: the mycobacterial proteome or the host cell proteome. The abundance of intracellular Zmp1 led to consider the hypothesis that it cleaves endogenous substrates. First, the proteome of BCG WT and Zmp1-mutant were compared by 2D-gel analysis. 21 spots recording an intensity ratio WT/mutant above 1.5 or below 0.6 have been identified. As expected by comparing different backgrounds, most of the hits corresponded to metabolic enzymes and housekeeping proteins. One hit, the product of the *M. tuberculosis* gene Rv2626c, was noteworthy for two reasons. First, it recorded the highest ratio WT/mutant in the assay and second, its biological function is closely correlated to mycobacterial virulence. The gene encodes for the hypoxic response protein 1 (Hrp1), a highly expressed member of the dormancy survival regulon DosR, and therefore involved in *M. tuberculosis* persistence. The protein is secreted from the bacterium in hypoxic *in vitro* culture

(Sharpe *et al.*, 2008). The small conserved protein (143 amino acids) activates macrophages by significant induction of pro-inflammatory cytokines (Bashir *et al.*, 2010). It is reasonable to assume that *M. tuberculosis* takes advantages from its degradation by the zinc metallopeptidase Zmp1 to dampen the host immune responses.

In order to reduce the variations derived by comparing two different genetic backgrounds, the proteome of BCG Δzmp1 was analyzed upon treatment with Zmp1. Only few spots showed a significant difference in intensity, demonstrating an absence of unspecific cleavage. The identified proteins were metabolic enzymes already found in the previous approach, and Rv2626c. In both approaches the protein spot Rv2626c appeared more intense in presence of Zmp1, and it recorded the highest ratio among the hits. The higher intensity of the spot in presence of the peptidase suggests that it could represent a cleavage product of the protein Rv2626c. Ideally, the 2D-gel analysis allows to detect the cleavage product and the corresponding full-length substrate on the same gel, as a proof of the proteolytic event. In the experiments described here, the full-length substrate was not detected in the control gel.

A major limitation of the 2D-PAGE is the inability to resolve extremely acidic, basic or hydrophobic proteins such as membrane-bound proteins, and also smaller proteins and peptides (<15 kDa). In addition, low abundance proteins are excluded as well, resulting overall in limited proteome coverage.

The 2D-PAGE technique, despite its limitations, provided an interesting clue in the study of endogenous substrates of Zmp1. The protein Rv2626c represents a putative Zmp1 substrate, and further investigation can corroborate this hypothesis. Moreover, considering the low proteome coverage of the method, other putative endogenous substrates might be envisaged. For instance, the small sized proteins have been excluded by the analysis, and they represent a likely target for Zmp1. Whether Zmp1 targets substrates within the host cell has to be investigated. As shown by Master and colleagues, Zmp1 might exert its activity out of the bacterial cell, within the phagosome environment or in the cytosol. Therefore the macrophage phagosome needs to be tested for searching Zmp1 proteolytic events. For such purpose, an

advanced proteomic techniques would be more suitable, such as mass spectrometry-based global proteome analysis.

## REFERENCES


Agard, N.J., and Wells, J.A. (2009) Methods for the proteomic identification of protease substrates. *Curr Opin Chem Biol* **13**: 503-509.

Bashir, N., Kounsar, F., Mukhopadhyay, S., and Hasnain, S.E. (2010) Mycobacterium tuberculosis conserved hypothetical protein rRv2626c modulates macrophage effector functions. *Immunology* **130**: 34-45.

Bredemeyer, A.J., Townsend, R.R., and Ley, T.J. (2005) Use of protease proteomics to discover granzyme B substrates. *Immunol Res* **32**: 143-153.

Ferraris, D.M., Sbardella, D., Petrera, A., Marini, S., Amstutz, B., Coletta, M., Sander, P., and Rizzi, M. (2011) Crystal structure of Mycobacterium tuberculosis zinc-dependent metalloprotease-1 (Zmp1), a metalloprotease involved in pathogenicity. *J Biol Chem* **286**: 32475-32482.

Hwang, I.K., Park, S.M., Kim, S.Y., and Lee, S.T. (2004) A proteomic approach to identify substrates of matrix metalloproteinase-14 in human plasma. *Biochim Biophys Acta* **1702**: 79-87.

Issaq, H., and Veenstra, T. (2008) Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): advances and perspectives. *Biotechniques* **44**: 697-698, 700.

Lee, A.Y., Goo Park, S., Kho, C.W., Young Park, S., Cho, S., Lee, S.C., Lee, D.H., Myung, P.K., and Park, B.C. (2004) Identification of the degradome of Isp-1, a major intracellular serine protease of Bacillus subtilis, by two-dimensional gel electrophoresis and matrix- assisted laser desorption/ionization-time of flight analysis. *Proteomics* **4**: 3437-3445.

Lopez-Otin, C., and Overall, C.M. (2002) Protease degradomics: a new challenge for proteomics. *Nat Rev Mol Cell Biol* **3**: 509-519.

Master, S.S., Rampini, S.K., Davis, A.S., Keller, C., Ehlers, S., Springer, B., Timmins, G.S., Sander, P., and Deretic, V. (2008) Mycobacterium tuberculosis prevents inflammasome activation. *Cell Host Microbe* **3**: 224-232.

Petrera, A., Amstutz, B., Gioia, M., Hähnlein, J., Baici, A., Selchow, P., Ferraris, D.M., Rizzi, M., Sbardella, D., Marini, S., Coletta, M., Sander, P. Functional characterization of the *Mycobacterium tuberculosis* zinc metallopeptidase Zmp1 and identification of potential substrates. *Biol Chem* **393**:631-640.

Sharpe, M.L., Gao, C., Kendall, S.L., Baker, E.N., and Lott, J.S. (2008) The structure and unusual protein chemistry of hypoxic response protein 1, a latency antigen and highly expressed member of the DosR regulon in Mycobacterium tuberculosis. *J Mol Biol* **383**: 822-836.

# 4. Functional characterization of the *Mycobacterium tuberculosis* zinc metallopeptidase Zmp1 and identification of potential substrates

Agnese Petrera*, Beat Amstutz*, Magda Gioia*, Janine Hähnlein, Antonio Baici, Petra Selchow, Davide M. Ferraris, Menico Rizzi, Diego Sbardella, Stefano Marini, Massimo Coletta and Peter Sander

*These authors contributed equally to this work

Biological Chemistry (2012) **393**, 631-640

**Abstract**

Zinc metallopeptidases of bacterial pathogens are widely distributed virulence factors and represent promising pharmacological targets. In this work we have characterized Zmp1, a zinc metallopeptidase identified as a virulence factor of *Mycobacterium tuberculosis* and belonging to the neprilysin (M13) family, whose X-ray structure has been recently solved. Interestingly, this enzyme shows an optimum activity toward a fluorogenic substrate at moderately acidic pH values (i.e., 6.3), which corresponds to those reported for the Mtb phagosome where this enzyme should exert its pathological activity. Substrate specificity of Zmp1 was investigated by screening a peptide library. Several sequences derived from biologically relevant proteins were identified as possible substrates, including the neuropeptides bradykinin, neurotensin and neuropeptide FF. Further, subsequences of other small bioactive peptides were found among most frequently cleaved sites, e.g. apelin-13 and substance P. We determined the specific cleavage site within neuropeptides by mass spectrometry, observing that hydrophobic amino acids, mainly phenylalanine and isoleucine, are overrepresented at position P1′. In addition, the enzymatic mechanism of Zmp1 toward these neuropeptides has been characterized, displaying some differences with respect to the synthetic fluorogenic substrate, and indicating that the enzyme adapts its enzymatic action to different substrates.

## Introduction

Tuberculosis (TB), the major cause of mortality around the world, kills about two million people annually, and approximately one-third of the world's population is asymptomatically infected with *Mycobacterium tuberculosis*, the main causative agent of the disease (WHO Fact sheet; http://www.who.int/mediacentre/factsheets/fs104/en/index.html). The prolonged coevolution of *M. tuberculosis* with its human hosts, and specifically with macrophages, has turned out in a peculiar survival strategy. Unlike other bacteria, which have developed different strategies, pathogenic mycobacteria (Mtb) have evolved to block lysosomal delivery (Nguyen and Pieters, 2005; Warner and Mizrahi, 2007). Hence, the pathogenic Mtb strategy demands that it actively subverts normal cellular mechanisms to avoid being killed (Nguyen and Pieters, 2005; Masjedi *et al.*, 2006).

The Rv0198c gene of Mtb encodes for the zinc metallopeptidase 1 (Zmp1) (663 aa), which, although not required for mycobacterial growth per se, turned out to be essential for survival inside host macrophages (Master *et al.*, 2008). An equivalent peptidase gene is also conserved in *Mycobacterium leprae* (MLCL622.12c; 667 aa), the causative agent of leprosy, whose genome exhibits massive gene decay (Cole *et al.*, 2001). Master and colleagues have described Zmp1 as a secreted virulence factor of *M. tuberculosis*, involved in phagosome maturation arrest. Zmp1 manipulates host immune defence by interfering with the activation of a multiprotein complex termed inflammasome (Master *et al.*, 2008). Furthermore, *zmp1* deletion influences the presentation of mycobacterial antigens, thus increasing the immunogenicity of BCG (Johansen *et al.*, 2011). The important role of Zmp1 in the Mtb pathogenicity has been recently confirmed in a mice model (Muttucumaru *et al.*, 2011), even though some questioning about the mechanism remains open.

Zmp1 belongs phylogenetically to the MEROPS peptidase family M13 and displays the conserved HExxH active site motif common to all $Zn^{2+}$ metallopeptidases. The targets of M13 members usually are bioactive peptides with a key function in the regulation of cardiovascular, nervous and immune systems (Turner *et al.*, 2001). Zmp1 shares particularly primary sequence similarity (48%)

with human neprilysin (749 aa), an extracellular peptidase that preferentially hydrolyzes extracellular oligopeptides (<5 kDa) on the amino side of hydrophobic residues (also including amyloid beta peptides implicated in Alzheimer's disease) (Carson and Turner, 2002; Madani *et al.*, 2006). We recently solved the crystal structure of Zmp1 and our investigation demonstrated that Zmp1 and neprilysin are also structurally conserved (Ferraris *et al.*, 2011). A major aspect concerns the active site of Zmp1, which is accessible only through a narrow channel, rendering difficult the interaction with large macromolecular substrates; therefore, the enzyme seems more tailored to process peptides. Here we report the characterization of the proteolytic activity of Zmp1 metallopeptidase, describing its substrate preferences, by employing a microarray based peptide library and analyzing its cleavage specificity by mass spectrometry. In addition, we have investigated the pH effect on Zmp1 catalysis, revealing an optimum activity at moderately acidic pH values which are substrate-dependent.

## Results

### Zmp1 activity depends on $Zn^{2+}$ binding

*M. tuberculosis zmp1* was cloned into a prokaryotic expression vector carrying a $His_6$ modification at its N-terminus and HA tag at its C-terminus for purification and detection. In addition, GST-*zmp1*-His/HA tagged protein was cloned. Both proteins were expressed to high levels in *E. coli* and purified in a two step approach using HisTrap column and ion exchange/GST-sepharose beads. Purity was > 95% according to a single band at the expected molecular weight in the SDS-PAGE (Figure 1A). Two mutants in the HExxH signature motif were produced by site-directed mutagenesis, namely H493A ($His^{493}$ was altered to Ala) and E494A mutant ($Glu^{494}$ was altered to Ala). Both mutated proteins were produced in a stable and soluble form and were purified to a high degree (Figures 1A and 1B).

In order to complement the crystallography data with the study of the protein in solution, circular dichroism spectroscopy was applied to Zmp1. In accord with the crystal structure (Ferraris *et al.*, 2011) all purified forms of Zmp1 (*i.e.*, wild type and the mutant) are properly folded, displaying identical CD spectra (Figure 1C). As predicted from the primary amino acid sequence and in agreement with

the crystal structure (Ferraris *et al.*, 2011), the secondary structure of Zmp1 displays over 70% of an α-helix structure, as estimated from the molar ellipticity at 222 nm (Morrow *et al.*, 2000). Salts, such as $CaCl_2$ and NaCl, do not alter the thermal stability of Zmp1 helical folding (data not shown). Zmp1 is able to cleave a synthetic fluorogenic substrate (see the materials and methods section), which is specific for a family of zinc endoproteinases, called matrix metalloproteinases (MMPs), sharing the catalytic motif (i.e., HExxH) with the M13 family. GST-Zmp1 wt cleaved the fluorogenic substrate with high efficiency and, as expected, mutating the catalytic motif almost abolished the proteolytic activity. Likewise, EDTA and the $Zn^{2+}$ chelator 1,10-phenanthroline reduced Zmp1 activity (see Table 1). Zmp1 activity was also strongly inhibited by the potent M13 metallopeptidases inhibitor phosphoramidon with a $K_I = 3.5(\pm 0.5)\times 10^{-8}$ M (Ferraris *et al.*, 2011). The enzyme does not undergo auto- proteolysis during the activity assay, as proved by the Selwyn's test (Selwyn, 1965) (see Figure S1).

**Table 1** Effect of peptidase inhibitors on the activity of Zmp1.

| Inhibitor | Concentration (mM) | Zmp1 activity (% of control) |
|---|---|---|
| Phosphoramidon | 0.1 | 5.1 ± 1.2 |
| 1,10-Phenanthroline | 0.1 | 7.8 ± 3.2 |
| EDTA | 1 | 46.4 ± 8.3 |

Zmp1 (10 nM) was preincubated with inhibitor for 10 min or 60 min (in case of EDTA) at room temperature. Activity toward the synthetic fluorogenic substrate was measured after 5 min at 37°C (±SD).

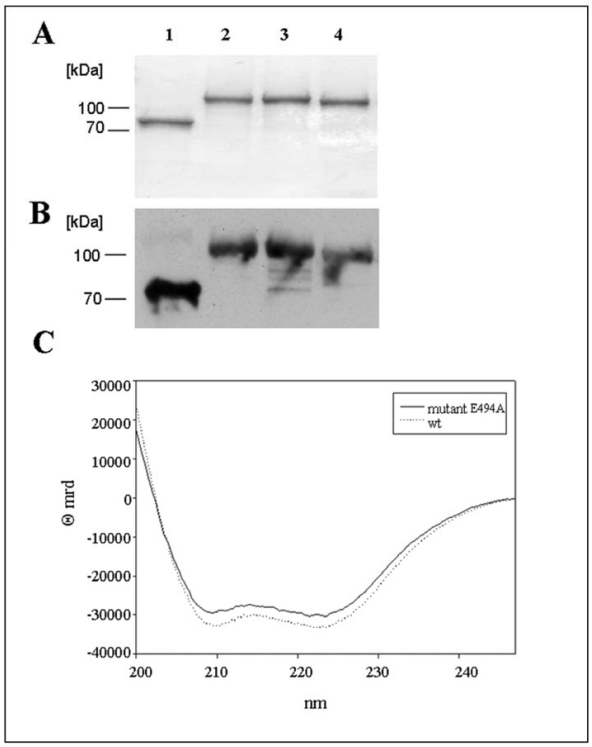



**Figure 1** Purification of Zmp1 and its structure in solution.
Coomassie stain of 12.5% SDS-Gel (A) and Western blotting (B) of Zmp1 variants. Lane 1: Zmp1 wt; Lane 2: GST-Zmp1; Lane 3: GST-Zmp1 E494A; Lane 4: GST-Zmp1 H493A. Zmp1 shows the predicted molecular mass of 72 kDa. The three GST-tagged variants have a molecular mass of approximately 100 kDa due to the fusion with a 26 kDa GST-tag. (C) Circular dichroism spectra of Zmp1. Dashed line corresponds to the circular dichroism spectrum of the wild type Zmp1, continuous line refers to the inactive E394A Zmp1.

**pH dependence of Zmp1 proteolytic activity**

The pH dependence of Zmp1 activity has been first measured employing a synthetic fluorogenic substrate. From the analysis of the dependence on the substrate concentration at each pH value according to Eqs (2) and (3) (see Materials and Methods) we have obtained the catalytic parameters (i.e., $k_{cat}/K_m$, $k_{cat}$ and $K_m$) at each investigated pH. Their pH dependence is reported in Figures 2A-2C for $k_{cat}/K_m$ (panel A), $k_{cat}$ (panel B) and $K_m$ (panel C), respectively, obtaining the $pK_a$ values reported in Table 2. Although the non-linear least-squares fitting of the pH dependence of individual parameters might have been carried out with only two protonatable groups, the global fitting of the pH

dependence of all three catalytic parameters demanded three protonatable groups [i.e., n = 3 in Eqs (4a-c), see materials and methods]. Evidently, the overall enzymatic activity of Zmp1 (corresponding to $k_{cat}/K_m$) is highest at pH 6.3 (Figure 2A), a pH value quite low compared to other metallopeptidases (Fasciglione *et al.*, 2000). Further, the pH dependence of $K_m$ (Figure 2C) indeed suggests that the substrate affinity tends to increase at pH lower than 7.5, since $K_m$ decreases. An interesting pattern is observed for the pH dependence of $k_{cat}$ (Figure 2B), where it appears evident that at least two groups with markedly different $pK_a$ values are modulating this parameter. In particular, it is worth underlining that the two-protonated species appears the most active form of Zmp1 and that this enhanced activity is connected to a residue characterized by a $pK_a = 6.73\pm0.17$ in the free enzyme, which then shifts to a $pK_a = 7.72\pm0.16$ in the substrate-bound form (Table 2). As a consequence, the pH-dependent profile of $k_{cat}$ appears somewhat displaced toward higher pH values.

**Table 2** Catalytic parameters for the different protonated species of Zmp1 toward the synthetic fluorogenic substrate and $pK_a$ values for the protonating groups in the free enzyme ($pK_U$) and of the substrate-bound enzyme ($pK_L$).

| | $k_{cat}/K_m$ ($M^{-1}$ $s^{-1}$) | $k_{cat}$ ($s^{-1}$) | $K_m$ (M) |
|---|---|---|---|
| 0-protonated | $5.0(\pm0.7) \times 10^3$ | 0.012±0.004 | $2.4(\pm0.4) \times 10^{-6}$ |
| 1-protonated | $7.1(\pm0.9) \times 10^3$ | 0.131±0.029 | $1.9(\pm0.3) \times 10^{-5}$ |
| 2-protonated | $3.0(\pm0.5) \times 10^4$ | 0.065±0.017 | $2.2(\pm0.5) \times 10^{-6}$ |
| 3-protonated | $9.5(\pm1.6) \times 10^2$ | 0.008±0.001 | $8.4(\pm1.6) \times 10^{-6}$ |

| | |
|---|---|
| $pK_{U1}$ | 8.17±0.16 |
| $pK_{U2}$ | 6.73±0.17 |
| $pK_{U3}$ | 5.43±0.19 |
| $pK_{L1}$ | 7.22±0.18 |
| $pK_{L2}$ | 7.72±0.17 |
| $pK_{L3}$ | 5.69±0.18 |

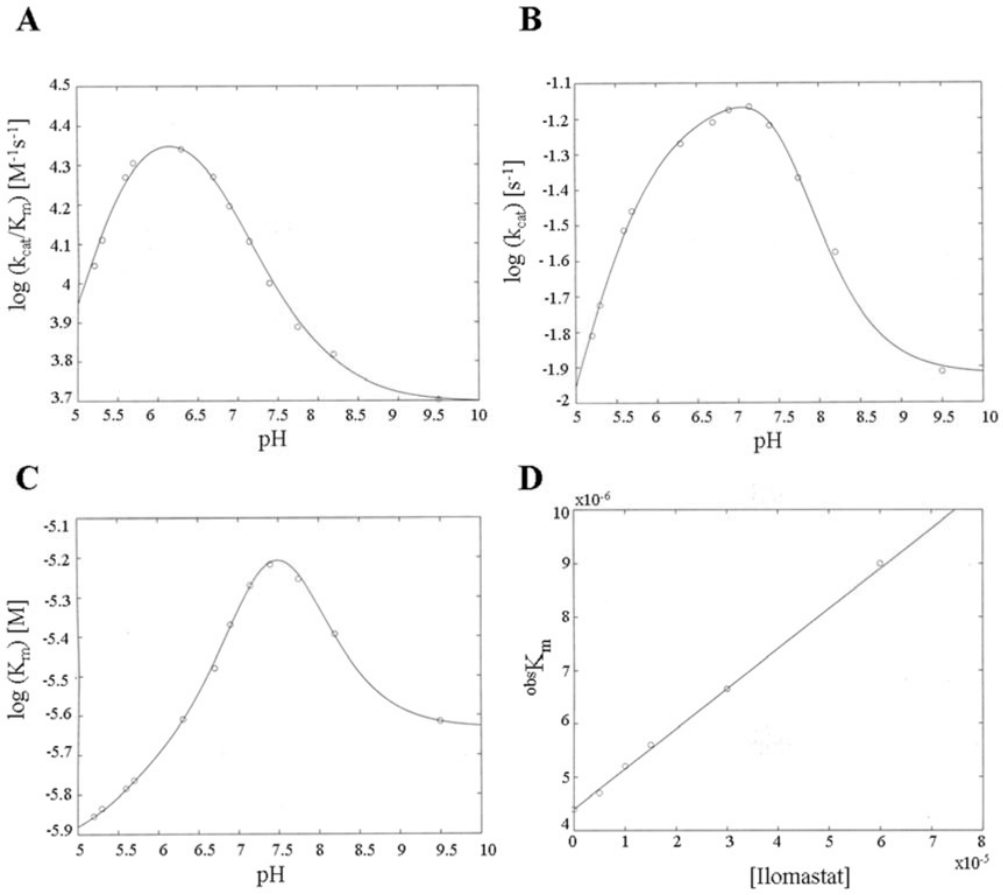



**Figure 2** pH dependence of catalytic parameters for Zmp1 at 37°C against a synthetic fluorogenic substrate. (A) pH dependence of $k_{cat}/K_m$. Continuous line represents the non-linear least squares fitting of data employing Eq. (4a) with parameters reported in Table 2. (B) pH dependence of $k_{cat}$. Continuous line represents the non-linear least squares fitting of data employing Eq. (4b) with parameters reported in Table 2. (C) pH dependence of $K_m$. Continuous line represents the non- linear least squares fitting of data employing Eq. (4c) with parameters reported in Table 2. (D) Ilomastat concentration dependence of substrate affinity constant $K_m$ for Zmp1 at 37°C. Values of $^{obs}K_m$ have been obtained by Lineweaver-Burk plot of omni-MMPs fluorogenic peptide cleavage in the presence of Zmp1 (0.4 nM) in absence and presence of different Ilomastat concentrations (namely 5, 10, 15, 30 and 60 μM). The continuous line is the non-linear least-squares fitting of data according to Eq. (5), from which a value of $^{0}K_m = 4.4(\pm0.6) \times 10^{-6}$ M and a value of $K_I = 5.9(\pm0.7) \times 10^{-5}$ are obtained.

**Zmp1 inhibition by Ilomastat**

The inhibition constant by Ilomastat has been determined for Zmp1, exploiting the inhibitor concentration dependence of the substrate affinity (Figure 2D). Application of Eq. (5) (see Materials and Methods) allows to determine the $K_I$ [= 5.9(±0.7)×$10^{-5}$ M], that is a value about $10^5$-fold higher than that observed for matrix metallopeptidases (Galardy *et al.*, 1994), reflecting a very low affinity of Ilomastat for Zmp1. This result indicates that Ilomastat is a very weak inhibitor for Zmp1, confirming the known difference between gluzincins and metzincins, as indeed it turns out to be evident also from the crystallographic structure, where the catalytic site results different from that of other metallopeptidases (Ferraris *et al.*, 2011).

**Zmp1 hydrolyzes a variety of bioactive peptides**

In order to possibly identify the native substrate(s) of Zmp1 we carried out an extensive screening of a microarray based peptide library. The library consists of 1989 octameric peptides derived from annotated cleavage sites of peptidases or random sequences and 1536 pentadecameric random peptide sequences. Peptides are C-terminally tagged with a phosphorylated tyrosine for detection. Peptide cleavage was quantified by incubation with anti-phosphotyrosine and fluorescent secondary antibody labeling and following fluorescence measurement. Fluorescence intensity of 51 out of 3525 peptide spots was decreased by more than 90% as compared to the BSA control. The 51 substrates were categorized into three groups (Table S1): 1) neuropeptides/substrates of NEP or ECE-1; 2) coagulation factors; 3) others. Interestingly, only short peptides (octameric) were represented among the top substrates. The relative frequency of each amino acid in the efficiently cleaved peptides was calculated by dividing the number of times this specific amino acid is represented in the screened substrates over the number of times that the amino acid is represented in the whole peptide microarray. This statistical analysis provides information about the amino acid preferences in the substrates, even though it does not allow prediction of actual cleavage sites. The most common motif among the top substrates was the sequence SPFR, which is present in the carboxy terminal part of the peptide hormone bradykinin. In particular, the statistical analysis showed an overrepresentation of arginine,

phenylalanine, isoleucine, leucine and valine in the top substrate sequences, whereas acidic amino acids like aspartic or glutamic acid were underrepresented, and no cysteine or tryptophan were present in any of the 51 top substrates.

**Zmp1 cleavage motif**

Several cleavage motifs can be described from the MALDI-TOF cleavage analysis of the investigated peptides (Figure 3A). The phenylalanine was the most represented residue at P1' position (Figure 3B), even though also isoleucine and leucine were abundant at this position. Proline was frequently found in the P3 site, although other positions near the cleavage site showed little preference. We conclude that PxxF is a preferred cleavage motif of Zmp1 and phenylalanine occupies P1' position.

**Identification of substrate cleavage sites**

First, we characterized the cleavage pattern of several neuropeptides identified in the peptide array (i.e., bradykinin, apelin-13, neurotensin, neuropeptide FF), including some NEP or ECE-1 substrates (i.e., substance P, dynorphin A (1-13) and angiotensin I), and the reaction by-products were analyzed by MALDI-TOF; in order to guarantee specific cleavage, high substrate/enzyme ratio was used (1000:1). Two cleavage sites were detected in bradykinin (Figure 4A), i.e., between $Pro^{7}$-$Phe^{8}$ and $Gly^{4}$-$Phe^{5}$, confirming the results obtained with the extended bradykinin, which is generated by the proteolytic cleavage of the precursor kininogen-1 by the enzyme kallikrein (Kaplan *et al.*, 2002). As a matter of fact, in rat, mouse, bovine and human the SPFR motif of kininogen-1 is conserved but it is followed by a different carboxy terminal amino acid. Four zinc metallopeptidases are mainly responsible for the metabolism of bradykinin, namely ACE, NEP, aminopeptidase P and carboxypeptidases M and N (Moreau *et al.*, 2005). Both cleavages resulted in two respective fragments that were detected by MALDI-TOF (Figures 4B and 4C).

We determined three cleavage sites in substance P: between $Gln^{6}$-$Phe^{7}$, $Phe^{7}$-$Phe^{8}$ and $Gly^{9}$-$Leu^{10}$ (Figure 4D). Interestingly, all the cleavage sites of bradykinin and substance P have a hydrophobic amino acid in the P1' site, mainly phenylalanine, which is typical of neprilysin, ECE-1 and thermolysin; further,

this observation is in agreement with the overrepresentation of this amino acid in the top substrates as compared to the total peptide library. Specific sites of cleavage within the other neuropeptides, such as angiotensin I (Figure S3), apelin-13 (Figure S4), dynorphin A (1 – 13) (Figure S5), and neurotensin (Figure S6) are summarized in Figure 3A.

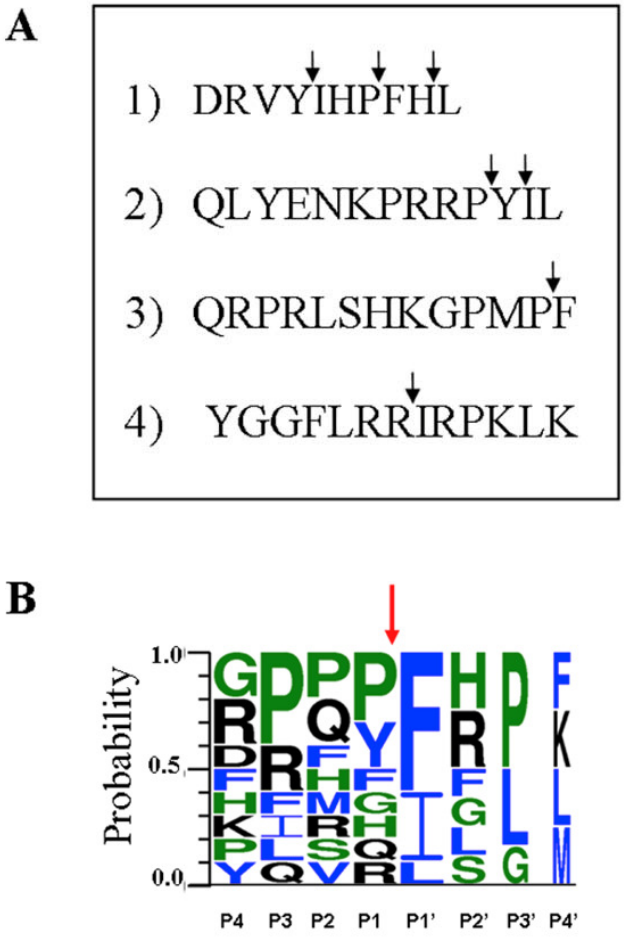



**Figure 3** Hydrolysis of hormone peptides by Zmp1. (A) Cleavage sites detected in the Zmp1 hydrolyzed fluorogenic peptides: 1) angiotensin I, 2) neurotensin, 3) apelin-13, 4) dynorphin A (1-13). Neuropeptide FF cleavage pattern was not studied because of poor detection signals in the MALDI-TOF analysis. (B) Amino acid preferences at various position of Zmp1 substrates calculated on the basis of the cleavage pattern of the studied hormone peptides. Size of amino acid letter is proportional to their occurrence in the recognition/cleavage motif. Cleavage site (marked with red arrow) is localized between position P1 and P1' (drawn using Weblogo 3.0).

Further, we synthesized a peptide fragment of mouse kininogen with the sequence **RPPGFSPFR**SVTV to map the cleavage site (bradykinin sequence in bold). The peptide was incubated with Zmp1 for 60 and 120 min, respectively, and products were analyzed by LC-MS. Cleavage sites were determined by calculating the molecular weight of putative products (data not shown). Primary cleavage occurred between $Pro^{7}$ and $Phe^{8}$ (Figure 4A, arrow with one star). This cleavage site lies within the bradykinin sequence and results in the inactivation of bradykinin rather than releasing bradykinin from its precursor. This peptide bond is also targeted by ACE, the main enzyme responsible for the metabolism of bradykinin (Blais *et al.*, 2000). Further cleavages were detected after longer incubation of the peptide with Zmp1, e.g. between $Gly^{4}$-$Phe^{5}$ and $Ser^{10}$-$Val^{11}$ (Figure 4A, arrows with two stars).

**Kinetic analysis of neuropeptides cleavage**

Zmp1 hydrolysis of the neuropeptides was further studied by the determination of kinetic constants $K_m$ and $k_{cat}$. Since the kinetic course of peptide cleavage by Zmp1 was affected by (a) product inhibition and/or (b) cleavage of multiple peptide bonds, initial velocities were exploited for calculating the kinetic parameters of the most susceptible peptide bond within a given peptide. Results are summarized in Table 3 as best fit values with associated standard errors from non-linear regression of the Michaelis-Menten equation fitted to data. The kinetic parameters for angiotensin I and neurotensin could not be calculated because even the initial velocity was strongly affected by the concurrent multiple cleavage and multiple product inhibition. Figure 5A shows, as an example, the double reciprocal Lineweaver-Burk plot for substance P, apelin-13 and bradykinin; it clearly emerges that substance P and apelin-13 are the best substrates among those investigated.

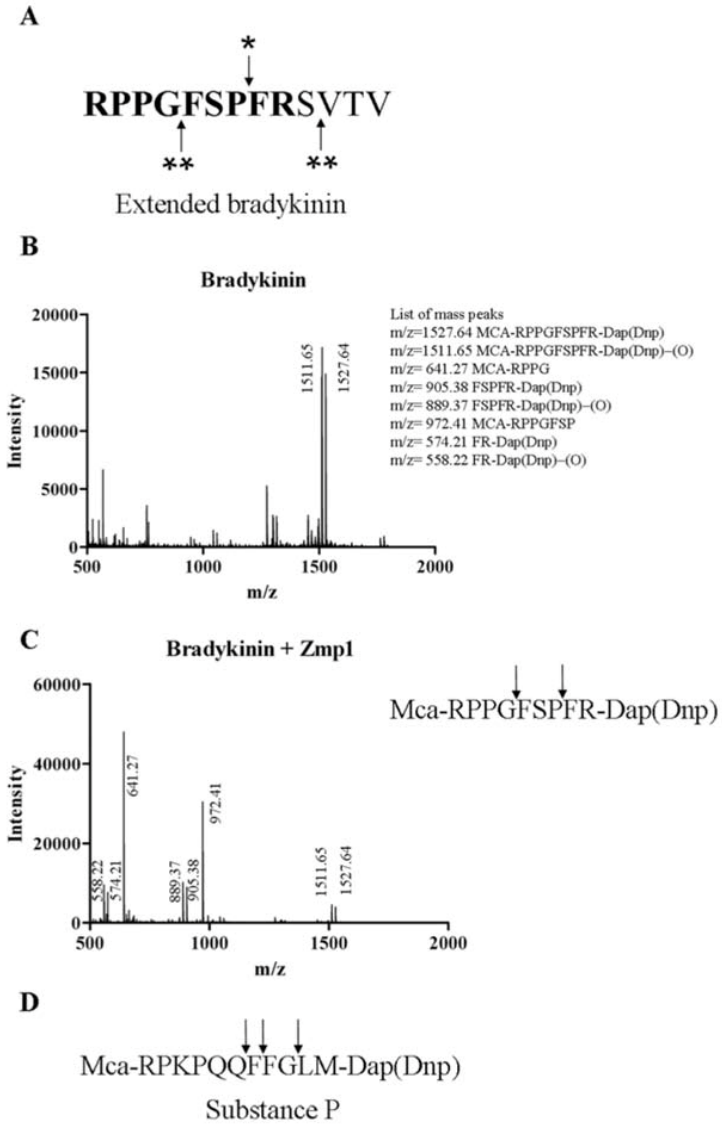



**Figure 4** Cleavage sites of bradykinin and substance P by Zmp1. (A) Determination of fragments of the 'extended' bradykinin (bradykinin sequence in bold) by Zmp1 through LC-MS analysis. Three different cleavage sites were detected in bradykinin. Primary site of hydrolysis is shown by the arrow with one asterisk, and secondary sites are shown by the arrows with two asterisks. (B) MALDI-TOF spectrum of fluorogenic bradykinin. (C) MALDI-TOF spectrum of Zmp1 hydrolyzed fluorogenic bradykinin. The C-terminal fragments showed an additional mass peak (Δ16 Da), due to a loss of oxygen (as described recently for nitro-phenyl-derivatives, see Petersson *et al.*, 2001). (D) Cleavage sites of fluorogenic substance P by Zmp1. The arrows above the amino acid sequence of substance P indicate the cleavage sites.

**Table 3** Kinetic parameters for the cleavage of the most susceptible peptide bond in five peptide substrates of Zmp1.

| Substrate | $k_{cat}$ ($s^{-1}$) | $K_m$ (μM) | $k_{cat}/K_m$ ($M^{-1}s^{-1}$) |
|---|---|---|---|
| Apelin | 1.54 ± 0.08 | 1.09 ± 0.23 | $1.41 \pm 0.31 \times 10^6$ |
| Substance P | 4.78 ± 0.71 | 3.92 ± 1.37 | $1.22 \pm 0.46 \times 10^6$ |
| Bradykinin | 2.76 ± 0.42 | 7.99 ± 1.45 | $3.46 \pm 0.67 \times 10^5$ |
| Dynorphin A | 2.25 ± 0.22 | 5.72 ± 1.22 | $3.93 \pm 0.92 \times 10^5$ |
| Neuropeptide FF | 1.07 ± 0.16 | 3.02 ± 1.15 | $3.56 \pm 1.45 \times 10^5$ |

Values of $k_{cat}$ and $K_m$, determined by non-linear regression from a fit of the Michaelis-Menten equation to primary data, are shown together with the standard errors from regression.

**pH dependence of substance P and apelin-13 hydrolysis**

In several studies ECE-1 was found to cleave neuropeptides inside acidified endosomes (i.e., at pH ≈ 5.5) (Padilla *et al.*, 2007; Roosterman *et al.*, 2007). We determined the pH dependence of Zmp1 activity for additional substrates, i.e., apelin-13 and substance P in order to see the pH preference for these substrates. Zmp1 cleaved the two neuropeptides preferentially in the pH range between 6.5 and 8.0 rather than at more acidic pH values, as observed for the MMP-2/MMP-7 substrate (Figure 2 for the pH profile of the MMP-2/MMP-7 substrate, Figure 5B for that of substance P and apelin-13). Data obtained from the pH-dependence of apelin-13 and substance P degradation by Zmp1 are reported in Table 4.

**Screening for macromolecular substrates (longer than 40 amino acids)**

Employing an in vitro proteolysis assay we have investigated the enzymatic action of Zmp1 on four of the substrates reported on the top substrates list of the peptide array, namely (i) β-amyloid, (ii) α-fibrinogen chain, (iii) insulin and (iv) denatured collagen. We have further tested Zmp1 proteolytic activity on caspase 1, the IL-1β activating enzyme. However, even prolonged exposition to Zmp1 (up to 24 hours) did not lead to any appreciable cleavage of these potential substrates (data not shown). In contrast, a control incubation with trypsin

led to complete digestion of the α-, β- and γ-fibrinogen chains already after 1 h of incubation.

**Table 4** Values of $k_{cat}/K_m$ ($M^{-1}s^{-1}$) for the different protonated species of Zmp1 toward apelin-13 and substance P and $pK_a$ values for the protonating groups in the free enzyme ($pK_U$).

| | Apelin-13 | Substance P |
|---|---|---|
| 0-protonated | $9.0(\pm1.6) \times 10^5$ | $8.8(\pm1.6) \times 10^5$ |
| 1-protonated | $1.7(\pm0.4) \times 10^6$ | $3.1(\pm0.7) \times 10^6$ |
| 2-protonated | $6.5(\pm1.1) \times 10^5$ | $1.6(\pm0.4) \times 10^5$ |
| 3-protonated | $1.0(\pm0.8) \times 10^4$ | $1.0(\pm0.7) \times 10^5$ |
| | | |
| $pK_{U1}$ | 8.17± 0.16 | |
| $pK_{U2}$ | 6.73±0.17 | |
| $pK_{U3}$ | 5.43±0.19 | |

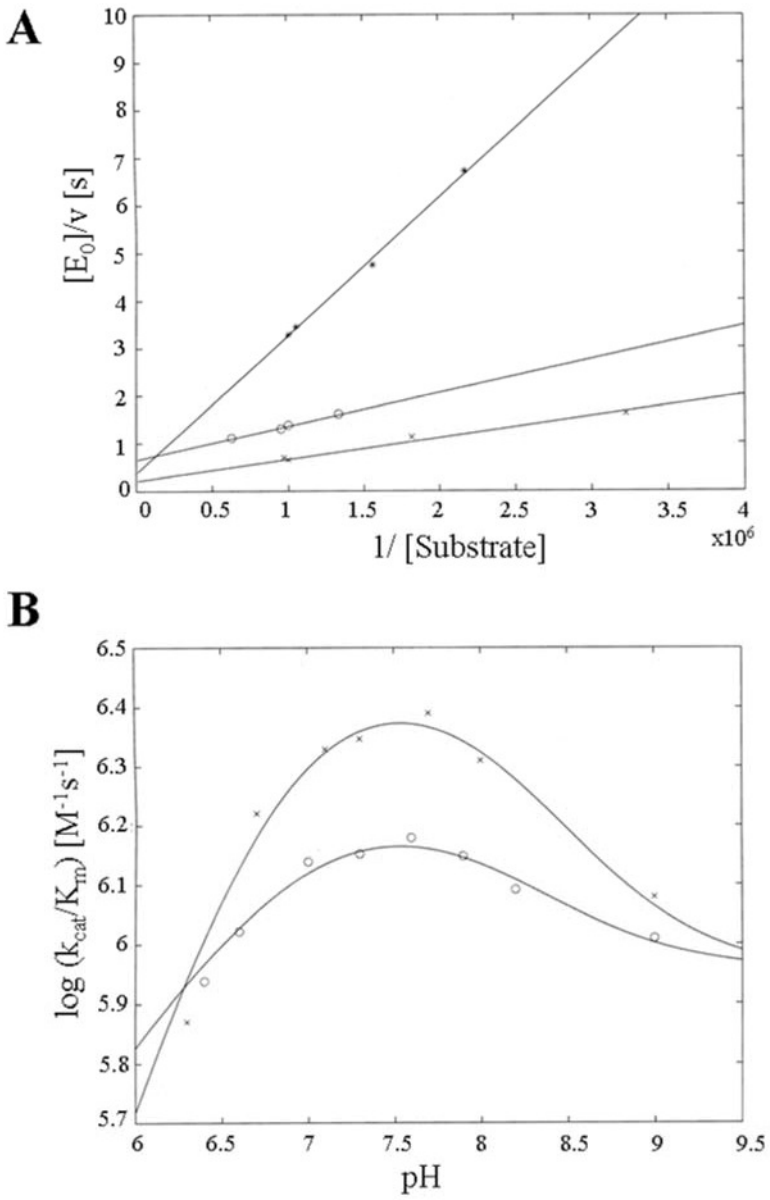



**Figure 5** Kinetic analysis of Zmp1 hydrolysis of apelin-13, substance P and bradykinin. (A) Double-reciprocal Lineweaver-Burk plot of the enzymatic processing by Zmp1 of apelin-13 (o), substance P (x) and bradykinin (*) at 37°C and pH 7.3. Solid lines represent the non- linear least-squares fitting of data according to Eq. (2). (B) pH dependence of the rate of cleavage ($k_{cat}/K_m$) of apelin-13 (o) and substance P (x) by Zmp1 at 37°C. Continuous lines represent the non-linear least-squares fitting of data according to Eq. (4a).

**Discussion**

Zinc metallopeptidases represent promising pharmacological targets and potential candidates for vaccine development (Miyoshi and Shinoda, 2000; Lopez-Otin and Bond, 2008; Baillie *et al.*, 2010). We have investigated the wild type and the inactive mutants (H493A and E494A) of the metallopeptidase Zmp1, which is an important virulence factor of *M. tuberculosis* (Master *et al.*, 2008).

Zmp1 belongs to the M13 family (MEROPS database http://merops.sanger.ac.uk /), like neprilysin (NEP) and endothelin-converting enzyme-1 (ECE-1), whose enzymatic action is restricted to peptides with a strong preference for cleaving the amino terminal bond of hydrophobic residues (Bland *et al.*, 2008). This preference for peptides is also confirmed by the recently solved X-ray crystallographic structure of Zmp1, which reveals that the active site is located at the bottom of a hydrophobic channel, likely impairing the access of large macromolecular substrates in the proximity of the reaction centre (Ferraris *et al.*, 2011).

In order to better characterize the specificity of Zmp1 a peptide microarray library was employed (including octameric and pentadecameric peptides); in particular, octameric peptides were enzymatically cleaved with the highest efficiency; further, a statistical analysis over the whole peptide array indicates a marked tendency of Zmp1 to cleave bonds where hydrophobic amino acids are located at the P1' position, similarly to NEP and ECE-1. In addition, as shown by the *in vitro* cleavage assay, sequences identified by the peptide array could not be cleaved when they were flanked by the whole protein sequences in a context of a structurally complex molecule (e.g. fibrinogen and β-amyloid). Among the hits screened several neuropeptides have been identified [i.e., substance P, neurotensin, bradykinin, apelin-13, angiotensin I, neuropeptide FF and dynorphin A (1-13)]. With respect to these, it is interesting to remark that a member of M13 family (ECE-1) localizes into the luminal side of endosomal membrane where it degrades internalized neuropeptides (e.g., substance P and apelin-13) (Roosterman *et al.*, 2007; El Messari *et al.*, 2004; Cottrell *et al.*, 2009), suggesting their possible biological relevance as substrates of this type of enzymes. As a matter of fact, besides the classical involvement of neuropeptides in neurotransmission and vasal smooth muscle cells constriction, at least for some of them, a novel role in

the modulation of immune responses has been envisaged (Skidgel *et al.*, 1984; Johnson *et al.*, 1999; Marriott and Bost, 2001; Yaraee *et al.*, 2007), affecting proinflammatory cytokine and chemokine activation (in the case of apelin-13) and macrophages activation (in the case of the substance P) in the early phases of immune responses to *Salmonella* infection (Kincy-Cain and Bost, 1996; O' Connor *et al.*, 2004; Leeper *et al.*, 2009).

A comparison between the degradation profile of these substrates by Zmp1, NEP and ECE-1 highlights overlapping specificities in terms of cleavage sites (i.e., for bradykinin, substance P and dynorphin A) with minor differences (i.e., secondary cleavage sites in angiotensin I and neurotensin) (Matsas *et al.*, 1984; Johnson *et al.*, 1999). However, a significant quantitative difference in terms of catalytic efficiency with respect to ECE-1 is observed with $k_{cat}/K_m$ ratios which in Zmp1 are 100- and 5-fold higher for substance P and bradykinin, respectively (Johnson *et al.*, 1999). On the whole, apelin-13 turns out to be an excellent substrate for Zmp1, being efficiently cleaved with the highest specificity constant (i.e., $k_{cat}/K_m$) among all tested peptides, a feature mostly related to the binding specificity rather than the cleavage efficiency.

Even though we cannot state at this stage whether these neuropeptides represent the Zmp1 substrate in vivo, pH scanning of Zmp1 proteolytic activity toward all investigated peptides indicates that Zmp1 can modulate its activity through the dynamical pH changes which occur in the maturating phagosome. In this respect, it is noteworthy that degradation of neuropeptides by ECE-1 in endosomes under acidic conditions (i.e., pH 5.5) is documented and it is associated to intracellular signaling and receptor turnover (El Messari *et al.*, 2004; Roostermann *et al.*, 2007; Cottrell *et al.*, 2009); analogously, we found that Zmp1 optimum of activity toward the synthetic substrate is at pH 6.3 (Figure 2A), a value corresponding to that of the early endosome (Pethe *et al.*, 2004). Therefore, it might be speculated that, similarly ECE-1, Zmp1 operates within the phagosome, arresting its maturation (Johansen *et al.*, 2011, Master *et al.*, 2008). As a matter of fact, at the acid pH of the maturating phagosome (i.e., pH 6.3) the substrate affinity of Zmp1 is fairly high (i.e., 2 μM), suggesting a conformation of the active site particularly suitable to interact with substrates under quite acidic pH conditions. In the case of

neuropeptides, such as apelin-13 and substance P the optimum pH is somewhat higher than that observed for the synthetic substrate, although $pK_a$ values for the pH-dependence of $k_{cat}/K_m$ are fully comparable. This strongly supports the idea that these $pK_a$ values refer only to the protonation of Zmp1 residues with no contribution from residues of the substrate(s). Therefore, since the whole enzymatic mechanism of Zmp1 seems to be modulated by the protonation of three classes of residues (characterized by $pK_{a1}$, $pK_{a2}$ and $pK_{a3}$, respectively), the different pH profile between these substrates can be simply explained by the involvement of different protonated species (the 2-protonated for the synthetic substrate and the 1-protonated for neuropeptides). This feature is clearly reflecting a relevant difference between the two classes of substrates in their ionic interactions with the enzyme.

These investigations demonstrate the importance of Zmp1 for mycobacteria host interaction and highlight its potential as a pharmacological target. Two major conclusions can be drawn, namely (i) Zmp1 preferentially cleaves small peptides, among which neuropeptides appear particularly interesting; (ii) Zmp1 adapts its activity to different microenvironmental conditions. The present biochemical characterization of the enzymatic properties of Zmp1 can be merged with the structural information on Zmp1 (Ferraris *et al.*, 2011), allowing to a structural-functional correlation, which will be of the utmost importance for a rational drug design.

## Materials and methods

### Molecular cloning and protein expression

*M. tuberculosis zmp1* gene was subcloned into pET100/D-TOPO® vector (Invitrogen) using a PCR amplified product from pMV361-*zmp1* vector, carrying *zmp1* gene, as previously reported (Master *et al.*, 2008). For all constructs, 1 L overnight culture of *E. coli* BL21 (Invitrogen) in LB medium (induced at $OD_{600}$ = 0.8 with 0.5 mM IPTG and grown at 18°C) was centrifuged at 5000 g for 20 min, followed by washing with PBS and additional centrifugation at 5000 g for 20 min. Cells were then lysed by three French press cycles (Thermo Electron) at $1.4 \times 10^8$

Pa, in the presence of peptidase inhibitors, followed by ultracentrifugation at 30000 g for 60 min. Supernatants were loaded on HisTrap™ HP column (GE Healthcare) equilibrated with buffer containing 20 mM $NaH_2PO_4$ pH 7.4, 0.5 M NaCl. Proteins were eluted with a concentration gradient of imidazole. Both Zmp1 and GST-Zmp1 were eluted at 250 mM imidazole (protein concentration was 0.5-1 mg/ml). After overnight dialysis against PBS, Zmp1 wt was further purified by ion exchange using a Q Column (GE Healthcare). GST-Zmp1 purification was performed using glutathione-sepharose beads (Glutathione Sepharose 4B, GE Healthcare). Proteins were dialyzed against PBS, and finally glycerol at a final concentration of 10% was added in order to store them in aliquots at -20°C. Proteins were analyzed by 12.5 % SDS-gel and subsequent Coomassie stain or Western Blot against HA.

**Circular dichroism experiments**

CD spectra for the validation of the correct folding of the protein were recorded on a Jasco J-710 spectropolarimeter equipped with a thermostated cell holder and connected to a data station for signal averaging and processing. All spectra were averages of four scans and were recorded employing quartz cells of 2-mm pathlength. Zmp1 was dissolved in CD buffer (20 mM Tris-HCl, 1 mM $CaCl_2$, 100mM NaCl) at about 2 µM concentration. The α-helix content of Zmp1 was calculated from the molar ellipticity of circular dichroic spectrum at 222 nm according to the following equation

$$\% \alpha\text{-helix} = (-[\theta]_{222nm} + 3000)/39000 \qquad \text{(Equation 1)}$$

where $[\theta]_{222nm}$ is the molar ellipticity at 222 nm, as previously described (Morrow *et al.*, 2000).

**Kinetic analysis and pH dependence of Zmp1 catalysis**

Internally quenched fluorogenic substrates with 7-methoxycoumarin (MCA) as fluorophore and 2,4-dinitrophenol (DNP) as quencher were either a synthetic fluorogenic substrate, displaying the following sequence: MCA[(7-methoxycoumarin-4-yl)acetyl]-Pro-Leu-Gly-Leu-DNP[*N*-3-(2,4-dinitrophenyl)-L-2,3-diaminopropionyl]-Ala-Arg-$NH_2$) or neuropeptides from EMC microcollections

(Tübingen, Germany). Neurotensin was synthesized with glutamine as N-terminal amino acid instead of pyroglutamic acid for ease of synthesis. The substrate peptides were prepared as stock solutions in 100% dimethyl sulfoxide (DMSO) at a final concentration of 1.0 mM and further diluted into the universal buffer (25 mM Bis-Tris-HCl, 25 mM Tris-HCl, 100 mM NaCl and 10 mM $CaCl_2$ prepared at 25°C) so as to maintain unaltered the composition at different pH values over the range investigated, keeping the DMSO concentration constant at 1% (v/v) for all dilutions. Substrate hydrolysis was measured fluorometrically with an Eclypse fluorimeter (Varian) or with a Flx800 spectrofluorometer (BIOTEK) equipped with KC Junior Software (BIOTEK) (λexc = 327 nm, λem = 393 nm); reactions were started by adding Zmp1 (at a final concentration of 10 nM) to a solution containing different substrates concentrations, spanning between 0.3 and 5 µM. For the pH dependence Zmp1 was incubated in the universal buffer between pH 5.2 and 9.5 for 5 min at 37°C; then substrate was added and fluorescence signal was followed. Background fluorescence of all buffers was similar. The initial velocities were measured at 37°C within a time interval during which the rate was constant and less than 10% of the substrate has been degraded. It ensured a steady-state condition and it was a prerequisite for the subsequent analysis step, which was based on the observation of an inverse linear correlation between velocity and substrate concentration according to Lineweaver-Burk equation (Eq.2) and Eadie-Hofstee linear regression (Eq.3)

$$\frac{E_0}{v} = \frac{K_m}{k_{cat}} \cdot \frac{1}{[S]} + \frac{1}{k_{cat}} \qquad \text{(Equation 2)}$$

$$\frac{v}{S} = \frac{V_{MAX}}{K_M} - \frac{v}{K_M} + \frac{K_M + S}{K_M S}\varepsilon \qquad \text{(Equation 3)}$$

where $E_0$ is the total enzyme concentration, v is the actual rate (expressed as mol/s), $K_m$ is the Michaelis-Menten equilibrium constant (expressed as mol), $k_{cat}$ is the rate-limiting step kinetic constant (expressed as $s^{-1}$), [S] is the substrate concentration and $V_{MAX}$ the ratio between $E_0$ and v. Linear regression plots were constructed from the velocity data and the catalytic parameters $k_{cat}$ and $K_m$ were extracted.

The pH dependence of the catalytic parameters has been fitted according to the following equations

$$ {}^{obs}k_{cat}/K_m = \frac{\sum_{i=0}^{i=n} {}^{i}k_{cat}/K_m \cdot \prod_{r=0}^{r=i} {}^{r}K_{ua} \cdot [H^+]^r}{\sum_{i=0}^{i=n} \prod_{r=0}^{r=i} {}^{r}K_{ua} \cdot [H^+]^r} \qquad \text{(Equation 4a)} $$

$$ {}^{obs}k_{cat} = \frac{\sum_{i=0}^{i=n} {}^{i}k_{cat} \cdot \prod_{r=0}^{r=i} {}^{r}K_{La} \cdot [H^+]^r}{\sum_{i=0}^{i=n} \prod_{r=0}^{r=i} {}^{r}K_{La} \cdot [H^+]^r} \qquad \text{(Equation 4b)} $$

$$ {}^{obs}K_m = {}^{0}K_m \cdot \frac{\sum_{i=0}^{i=n} \prod_{r=0}^{r=i} {}^{r}K_{ua} \cdot [H^+]^r}{\sum_{i=0}^{i=n} \prod_{r=0}^{r=i} {}^{r}K_{La} \cdot [H^+]^r} \qquad \text{(Equation 4c)} $$

where ${}^{obs}k_{cat}/K_m$, ${}^{obs}k_{cat}$ and ${}^{obs}K_m$ are the observed parameter(s) as a function of pH, ${}^{i}k_{cat}/K_m$, ${}^{i}k_{cat}$ and ${}^{i}K_m$ are the parameter(s) for the $i^{th}$-protonated form(s), ${}^{i}K_{ua}$ are the $i^{th}$ protonation constant(s) for the free enzyme E and ${}^{i}K_{La}$ are the $i^{th}$ protonation constant(s) for the substrate- bound enzyme ES when it is involved in the cleavage event at the rate-limiting step. The global fitting procedure, employing simultaneously Eqs. (4a-4c) to describe the pH dependence of all three catalytic parameters of Zmp1 (i.e., $k_{cat}/K_m$, $k_{cat}$ and $K_m$), allows to determine the $pK_a$ of protonating groups, which modulate the enzymatic activity of Zmp1.

**Inhibition assay**

For inhibition experiments, Zmp1 (10 nM) was preincubated with EDTA, phosphoramidon and 1,10-phenanthroline (Sigma) in the universal buffer (see above) at RT before addition of the synthetic fluorogenic substrate. Fluorescence was measured after 5 min of incubation with the fluorogenic substrate (5 µM) at 37°C. In the case of Ilomastat (British Biotech Pharmaceutical, Cowley, Oxford, UK) the inhibition constant $K_I$ was measured obtaining the various $K_m$ from the substrate concentration dependence at various inhibitor concentration, employing

Eqs (2) and (3) and then plotting data of $K_m$ as a function of inhibitor concentration, as from the following equation

$$K_m = {}^0K_m \cdot (1 + \frac{[I]}{K_I}) \qquad \text{(Equation 5)}$$

where $K_m$ is the Michaelis constant at different inhibitor concentrations, ${}^0K_m$ is the Michaelis constant in the absence of the inhibitor, [I] is the inhibitor concentration and $K_I$ is the inhibition constant.

**Peptide array**

Peptide Array was performed at JPT Berlin using a commercially available microarray (JPT Peptide Technologies GmbH, Berlin, Germany; batch# 1562, both for peptidase incubation and control incubation in triplicates). The library consists of 1989 octameric peptides derived from annotated cleavage sites of peptidases or random sequences and 1536 pentadecameric random peptide sequences. Peptides contain a C-terminal phosphorylated tyrosine for detection with an anti-phosphotyrosine antibody. The peptide microarray was incubated with 10 µg/ml Zmp1 (total reaction volume 400 µl) for 4 h at 37°C, using a microarray sandwich like construction. As a control sample an additional incubation was carried out in the absence of the target peptidase, which was followed by washing steps with 50 mM TBS-buffer, pH 7.2 and dd$H_2O$. Subsequent drying with a microarray centrifuge was performed. Microarrays were first incubated with anti-pTyr-antibody (for about 45 minutes) and then for 30 minutes with Dylight 649-labelled anti-mouse-antibody. For detection, a Tecan HS4800 microarray processing station was employed. Total incubation time was 45 minutes for anti-pTyr-100- antibody and 30 minutes for anti-mouse-antibody, followed by washing steps with 50 mM TBS pH 7.2 and SSC buffer, pH 7.0 (JPT). Finally, the microarray plates were dried using a nitrogen stream. Analysis was performed with high resolution fluorescence scanner (Axon Genepix 4200 AL Scanner) and Spot-recognition software Genepix 6.0 and Microsoft Excel was employed for data analysis. A reduction by more than 90% of signal intensity was considered as effective cleavage.

**Cleavage site determination by MALDI-TOF**

Peptide cleavage assay was performed as following: 100 µM neuropeptides were incubated with or without Zmp1 (20 nM) for 5 min at 37°C in a total volume of 200 µl. Then, 1 µl of a 1:100 dilution with MALDI Matrix solution (consisting of 60% acetonitrile, 0.1% TFA in dd$H_2O$ and 4 mg/ml α-cyano-4-hydroxy cinnamic acid) was spotted the MALDI plate. After evaporation of solutes, MALDI spectra were recorded with a 4800 plus MALDI TOF/TOF Analyzer (Applied Biosystems, Foster City, CA). MALDI was equipped with a 355 nm Nd:YAG laser which operates at a repetition rate of 200 Hz. MS spectra were acquired in the reflectron mode, typically in the mass range from 600 to 4000 Da. The MS spectra were externally calibrated using a standard peptide mixture (Applied Biosystems). Recording software was 4000 Series Explorer and for the final preparation of the figures, Data Explorer, Microsoft Excel and Prism Graph were used.

**LC-MS of bradykinin sequence**

Bradykinin-elongated peptide with the sequence $H_2N$-RPPGFSPFRSVTV-$NH_2$ (C-terminally tagged with an amide group to avoid negative charge effects) was synthesized by JPT and dissolved in the assay buffer (100 mM Tris pH 7.3, 100 mM NaCl, 10 mM $CaCl_2$) at a final concentration of 1 mg/ml. Zmp1 was added to individual vials at concentrations of 10, 40 and 200 nM; a vial was kept in the absence of peptidase as a control. At different time intervals 8 µl were injected into LC-MS system and UV traces were recorded with Agilent LC-1200 VWD Detector at 220 nm; MS traces were recorded using Agilent Single-Quad G6130 mass spectrometer. All evaluations were performed with Agilent chemstation B0301 data analysis software.

**Acknowledgments**

We acknowledge the support from the Functional Genomics Center Zurich (FGCZ). This work was also supported in part by the Swiss National Science Foundation (3100A0-135705 to P.S.), the Italian Ministry for University and Research (PRIN 200993WWF9 to M.C.) and the European Union (EU-PF7 New TB Vac, project no. 241745 and EU-PF7 SysteMtb Collaborative Project no. 241587). B.A. was supported by the Forschungskredit 2010 from University of Zurich (no. 54232101).

**References**

Baillie, L. W., Huwar, T. B., Moore, S., Mellado-Sanchez, G., Rodriguez, L., Neeson, B. N., Flick-Smith, H. C., Jenner, D. C., Atkins, H. S., Ingram, R. J., Altmann, D. M., Nataro, J. P., and Pasetti, M. F. (2010) An anthrax subunit vaccine candidate based on protective regions of Bacillus anthracis protective antigen and lethal factor. *Vaccine* 28, 6740-6748.

Blais, C., Jr., Marceau, F., Rouleau, J. L., and Adam, A. (2000) The kallikrein-kininogen- kinin system: lessons from the quantification of endogenous kinins. *Peptides* **21**, 1903- 1940.

Bland, N. D., Pinney, J. W., Thomas, J. E., Turner, A. J., and Isaac, R. E. (2008) Bioinformatic analysis of the neprilysin (M13) family of peptidases reveals complex evolutionary and functional relationships. *BMC Evol. Biol.* **8**, 16.

Carson, J. A., and Turner, A. J. (2002) Beta-amyloid catabolism: roles for neprilysin (NEP) and other metallopeptidases? *J. Neurochem.* **81**, 1-8.

Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honore, N., Garnier, T., Churcher, C., Harris, D., *et al.* (2001) Massive gene decay in the leprosy bacillus. Nature 409, 1007-1011.

Cottrell, G. S., Padilla, B. E., Amadesi, S., Poole, D. P., Murphy, J. E., Hardt, M., Roosterman, D., Steinhoff, M., and Bunnett, N. W. (2009)

Endosomal endothelin- converting enzyme-1: a regulator of beta-arrestin-dependent ERK signaling. *J. Biol. Chem.* **284**, 22411-22425.

El Messari, S., Iturrioz, X., Fassot, C., De Mota, N., Roesch, D., and Llorens-Cortes, C. (2004) Functional dissociation of apelin receptor signaling and endocytosis: implications for the effects of apelin on arterial blood pressure. *J. Neurochem.* **90**, 1290- 1301.

Fasciglione, G. F., Marini, S., D'Alessio, S., Politi, V., and Coletta, M. (2000) pH- and temperature-dependence of functional modulation in metalloproteinases. A comparison between neutrophil collagenase and gelatinases A and B. *Biophys. J.* **79**, 2138-2149.

Ferraris, D. M., Sbardella, D., Petrera, A., Marini, S., Amstutz, B., Coletta, M., Sander, P., and Rizzi, M. (2011) Crystal structure of Mycobacterium tuberculosis zinc-dependent metalloprotease-1 (Zmp1), a metalloprotease involved in pathogenicity. *J. Biol. Chem.* **286**, 32475-32482.

Galardy, R. E., Cassabonne, M. E., Giese, C., Gilbert, J. H., Lapierre, F., Lopez, H., Schaefer, M. E., Stack, R., Sullivan, M., Summers, B. and *et al.* (1994) Low molecular weight inhibitors in corneal ulceration. *Ann. N. Y. Acad. Sci.* **732**, 315-323.

Johansen, P., Fettelschoss, A., Amstutz, B., Selchow, P., Waeckerle-Men, Y., Keller, P., Deretic, V., Held, L., Kundig, T. M., Bottger, E. C., and Sander, P. (2011) Relief from Zmp1-mediated arrest of phagosome maturation is associated with facilitated presentation and enhanced immunogenicity of mycobacterial antigens. *Clin. Vaccine Immunol.* **18**, 907-913.

Johnson, G. D., Stevenson, T., and Ahn, K. (1999) Hydrolysis of peptide hormones by endothelin-converting enzyme-1. A comparison with neprilysin. *J. Biol. Chem.* **274**, 4053-4058.

Kaplan, A. P., Joseph, K., and Silverberg, M. (2002) Pathways for bradykinin formation and inflammatory disease. *J. Allergy Clin. Immunol.* **109**, 195-209.

Kincy-Cain, T., and Bost, K. L. (1996) Increased susceptibility of mice to Salmonella infection following in vivo treatment with the substance P antagonist, spantide II. *J. Immunol.* **157**, 255-264.

Leeper, N. J., Tedesco, M. M., Kojima, Y., Schultz, G. M., Kundu, R. K., Ashley, E. A., Tsao, P. S., Dalman, R. L., and Quertermous, T. (2009) Apelin prevents aortic aneurysm formation by inhibiting macrophage inflammation. *Am. J. Physiol. Heart Circ. Physiol.* **296**, H1329-H1335.

Lopez-Otin, C., and Bond, J. S. (2008) Proteases: multifunctional enzymes in life and disease. *J. Biol. Chem.* **283**, 30433-30437.

Madani, R., Poirier, R., Wolfer, D. P., Welzl, H., Groscurth, P., Lipp, H. P., Lu, B., El Mouedden, M., Mercken, M., Nitsch, R. M., and Mohajeri, M. H. (2006) Lack of neprilysin suffices to generate murine amyloid-like deposits in the brain and behavioral deficit in vivo. *J. Neurosci. Res.* **84**, 1871-1878.

Marriott, I., and Bost, K. L. (2001) Substance P receptor mediated macrophage responses. *Adv. Exp. Med. Biol.* **493**, 247-254.

Masjedi, M. R., Farnia, P., Sorooch, S., Pooramiri, M. V., Mansoori, S. D., Zarifi, A. Z., Akbarvelayati, A., and Hoffner, S. (2006) Extensively drug-resistant tuberculosis: 2 years of surveillance in Iran. *Clin. Infect. Dis.* **43**, 841-847.

Master, S. S., Rampini, S. K., Davis, A. S., Keller, C., Ehlers, S., Springer, B., Timmins, G. S., Sander, P., and Deretic, V. (2008) Mycobacterium tuberculosis prevents inflammasome activation. *Cell Host Microbe* **3**, 224-232.

Matsas, R., Kenny, A. J., and Turner, A. J. (1984) The metabolism of neuropeptides. The hydrolysis of peptides, including enkephalins, tachykinins and their analogues, by endopeptidase-24.11. *Biochem. J.* **223**, 433-440.

Miyoshi, S., and Shinoda, S. (2000) Microbial metalloproteases and pathogenesis. *Microbes Infect.* **2**, 91-98.

Moreau, M. E., Garbacki, N., Molinaro, G., Brown, N. J., Marceau, F., and Adam, A. (2005) The kallikrein-kinin system: current and future pharmacological targets. *J. Pharmacol. Sci.* **99**, 6-38.

Morrow, J. A., Segall, M. L., Lund-Katz, S., Phillips, M. C., Knapp, M., Rupp, B., and Weisgraber, K. H. (2000) Differences in stability among the human apolipoprotein E isoforms determined by the amino-terminal domain. *Biochemistry* **39**, 11657-11666.

Muttucumaru, D.G., Smith, D.A., McMinn, E.J., Reese, V., Coler, R.N., and Parish, T. (2011) Mycobacterium tunerculosis Rv0198c, a putative matrix metalloprotease, is involved in pathogenicity. *Tuberculosis* (Edinb) **91**, 111-116.

Nguyen, L., and Pieters, J. (2005) The Trojan horse: survival tactics of pathogenic mycobacteria in macrophages. *Trends Cell Biol.* **15**, 269-276.

O'Connor, T. M., O'Connell, J., O'Brien, D. I., Goode, T., Bredin, C. P., and Shanahan, F. (2004) The role of substance P in inflammatory disease. *J. Cell Physiol.* **201**, 167-180.

Padilla, B. E., Cottrell, G. S., Roosterman, D., Pikios, S., Muller, L., Steinhoff, M., and Bunnett, N. W. (2007) Endothelin-converting enzyme-1 regulates endosomal sorting of calcitonin receptor-like receptor and beta-arrestins. *J. Cell Biol.* **179**, 981-997.

Petersson, A. S., Steen, H., Kalume, D. E., Caidahl, K., and Roepstorff, P. (2001) Investigation of tyrosine nitration in proteins by mass spectrometry. *J. Mass Spectrom.* **36**, 616-625.

Pethe, K., Swenson, D. L., Alonso, S., Anderson, J., Wang, C., and Russell, D. G. (2004) Isolation of Mycobacterium tuberculosis mutants defective in the arrest of phagosome maturation. *Proc. Natl. Acad. Sci. USA* **101**, 13642-13647.

Roosterman, D., Cottrell, G. S., Padilla, B. E., Muller, L., Eckman, C. B., Bunnett, N. W., and Steinhoff, M. (2007) Endothelin-converting enzyme 1 degrades neuropeptides in endosomes to control receptor recycling. *Proc. Natl. Acad.*

*Sci. USA* **104**, 11838-11843.

Selwyn, M. J. (1965) A simple test for inactivation of an enzyme during assay. *Biochim. Biophys. Acta* **105**, 193-195.

Skidgel, R. A., Engelbrecht, S., Johnson, A. R., and Erdos, E. G. (1984) Hydrolysis of substance p and neurotensin by converting enzyme and neutral endopeptidase. *Peptides* **5**, 769-776.

Turner, A. J., Isaac, R. E., and Coates, D. (2001) The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function. *Bioessays* **23**, 261-269.

Warner, D. F., and Mizrahi, V. (2007) The survival kit of Mycobacterium tuberculosis. *Nat. Med.* **13**, 282-284

Yaraee, R., Ebtekar, M., Ahmadiani, A., Sabahi, F., and Ghazanfari, T. (2007) The effect of substance P on nitric oxide production by HSV-1 infected macrophages. *Int. Immunopharmacol.* **7**, 135-139.

## Supplementary data

### Table S1. List of protein subsequences efficiently cleaved by Zmp1

Depicted is the ranking of particular cleavage sequences according to the cleavage rate. The name of the protein, the annotation, the specific sequence and the functional group are given.

| Top substrate | Protein | Annotation | Amino acid sequence | Cleavage rate [%] | Group |
|---|---|---|---|---|---|
| 1 | random | detected by MS | TRQFFESF | 100 | 3 |
| 2 | Collagen | CA13_HUMAN | GPLGIAGI | 98 | 3 |
| 3 | random | detected by MS | VSTQFIRA | 98 | 3 |
| 4 | Chymotrypsinogen B | CTRB_HUMAN | GLSRIVNG | 98 | 3 |
| 5 | Lantibiotic nisin-A | LANN_LACLA | ASPRITSI | 98 | 3 |
| 6 | Matrix metalloprotease 9 | MM09_HUMAN | DLGRFQTF | 98 | 3 |
| 7 | Apolipoprotein AI | APA1_HUMAN | LESFKVSF | 97 | 3 |
| 8 | Putative protein | C0AAH1 | NIILSLIM | 97 | 3 |
| 9 | Chymoptrypsin like elastase | EL2A_HUMAN | YVTRVVGG | 96 | 3 |
| 10 | Mellitin | MEL_POLHE | GIGAVLKV | 96 | 3 |
| 11 | Beta-endorphine | COLI_HUMAN | TLFKNAII | 96 | 1 |
| 12 | Collagenase | MM01_BOVIN | GIQAIYGP | 96 | 3 |
| 13 | Actin | ACTA_HUMAN | GILTLKYP | 95 | 3 |
| 14 | T-kininogen 1 rat | KNT1_RAT | SPFRLVRV | 95 | 1 |
| 15 | Bovine casein | CAS1_BOVIN | LLRFFVAP | 95 | 3 |
| 16 | Coagulation factor V | FA5_HUMAN | LGIRSFRN | 94 | 2 |
| 17 | Kininogen-1 mouse | KNG_MOUSE | SPFRSVTV | 94 | 1 |
| 18 | Lassa virus glycoprotein | VGLY_LASSG | RRLLGTFT | 94 | 3 |
| 19 | Coagulogen | COAG_TACTR | VSGRGFSI | 93 | 2 |
| 20 | Kallikrein-1 human | KLK1_HUMAN | IQSRIVGG | 93 | 2 |
| 21 | Proclotting enzyme | PCE_TACTR | TTTRIIGG | 93 | 2 |
| 22 | Neurotensin | NEUT_HUMAN | NKPRRPYI | 93 | 1 |
| 23 | Apelin | APEL_HUMAN | GRRKFRRQ | 93 | 1 |
| 24 | Coagulation factor II (thrombin) receptor | Q3UK34 | VNPRSFFL | 93 | 2 |
| 25 | Atrial natriuretic factor | ANF_HUMAN | RALLTAPR | 93 | 1 |
| 26 | MAPKK2 | MPK2_HUMAN | PVLPALTI | 93 | 3 |
| 27 | random | detected by MS | VVVYSFTA | 93 | 3 |
| 28 | random | detected by MS | RQLYFFHS | 93 | 3 |
| 29 | Alpha-2 antiplasmin | A2AP_HUMAN | QVSPLTLL | 93 | 2 |
| 30 | Hatching enzyme of common sea urchin | HE_PARLI | GIRSLYGS | 93 | 3 |
| 31 | Fibrinogen alpha chain | FIBA_HUMAN | SSSYSKQF | 93 | 2 |
| 32 | Apelin | APEL_HUMAN | NVRHLVQP | 93 | 1 |
| 33 | random | detected by MS | IYQQFNYT | 92 | 3 |
| 34 | random | detected by MS | VIEFRVMV | 92 | 3 |
| 35 | Complement factor B | CFAB_HUMAN | QQKRKIVL | 92 | 3 |
| 36 | Transcription factor ATF-6 beta | AT6B_HUMAN | RHLLGFSE | 92 | 3 |
| 37 | Urokinase-type plasminogen activator | UROK_HUMAN | LRPRFKII | 92 | 2 |
| 38 | Hemoglobin subunit B | HBG_HUMAN | GRLLVVYP | 92 | 3 |
| 39 | Tobacco virus polyprotein | POLG_TVMV | RARFSQSI | 92 | 3 |
| 40 | Coagulation factor X | FA10_HUMAN | NLTRIVGG | 92 | 2 |
| 41 | Bovine casein | CASB_BOVIN | PIQAFLLY | 92 | 3 |
| 42 | Human insulin | INS_HUMAN | GAGSLQPL | 92 | 1 |
| 43 | random | detected by MS | GVEFRVAA | 91 | 3 |
| 44 | Amyloid beta protein | A4_HUMAN | GGVVIATV | 91 | 1 |
| 45 | Human insulin | INS_HUMAN | HLVEALYL | 91 | 1 |
| 46 | random | detected by MS | PHASFIAI | 91 | 3 |
| 47 | Complement C1r subcomponent | C1R_HUMAN | QRQRIIGG | 91 | 3 |
| 48 | Alpha-1 antichymotrypsin | AACT_HUMAN | LLSALVET | 91 | 3 |
| 49 | Kininogen-1 bovine | KNH1_BOVIN | SPFRSVQV | 91 | 1 |
| 50 | Yeast heat shock glycoprotein | CCW7_YEAST | AKRAASQI | 91 | 3 |
| 51 | Neuropeptide FF | NPFF_HUMAN | SQAFLFQP | 91 | 1 |

**Figure S1. Selwyn stability test on Zmp1**

RFU plotted as a function of the product of enzyme concentration and time. 50, 100, 150 nM Zmp1 were incubated with synthetic substrate over a period of 200 s. The points lie close together within the limits of experimental error for different enzyme concentrations.

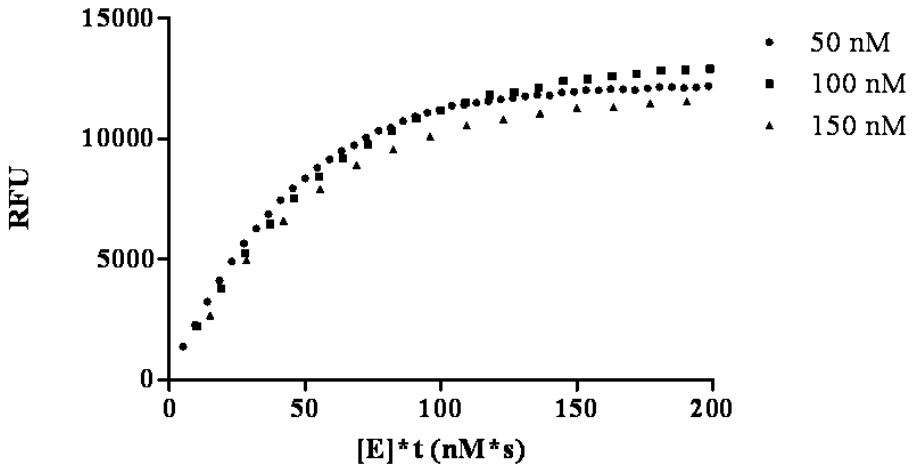


**Figure S2. Cleavage sites of substance P by Zmp1**

Determination of fragments of substance P by Zmp1 through LC-MS analysis. (A) MALDI-TOF spectrum of fluorogenic substance P. (B) MALDI-TOF spectrum of Zmp1 hydrolyzed fluorogenic substance P. The arrows above the amino acid sequence of substance P indicate the cleavage sites.

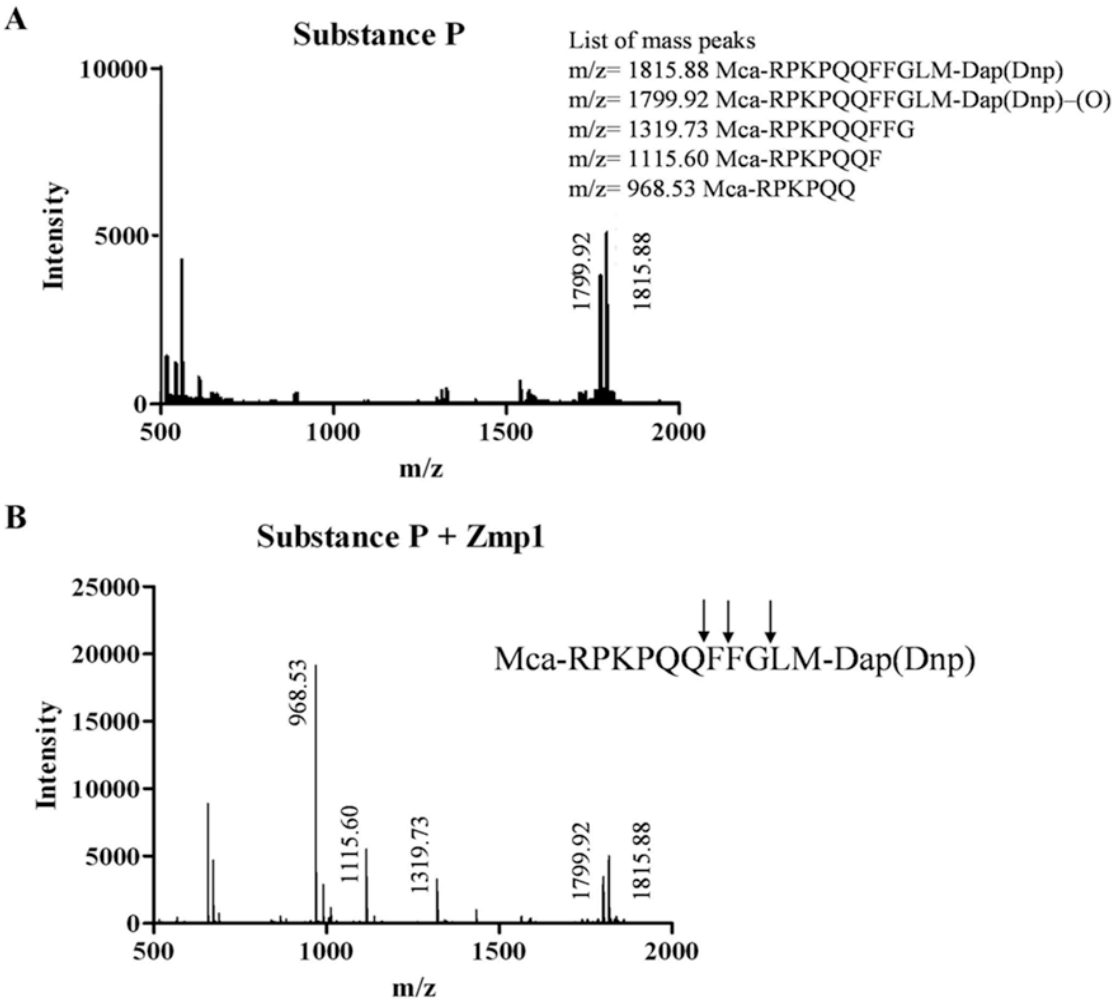


**Figure S3. Cleavage sites of angiotensin I by Zmp1**

Determination of fragments of angiotensin I by Zmp1 through LC-MS analysis. (A) MALDI-TOF spectrum of fluorogenic angiotensin I. (B) MALDI-TOF spectrum of Zmp1 hydrolyzed fluorogenic angiotensin I. The arrows above the amino acid sequence of angiotensin I indicate the cleavage sites

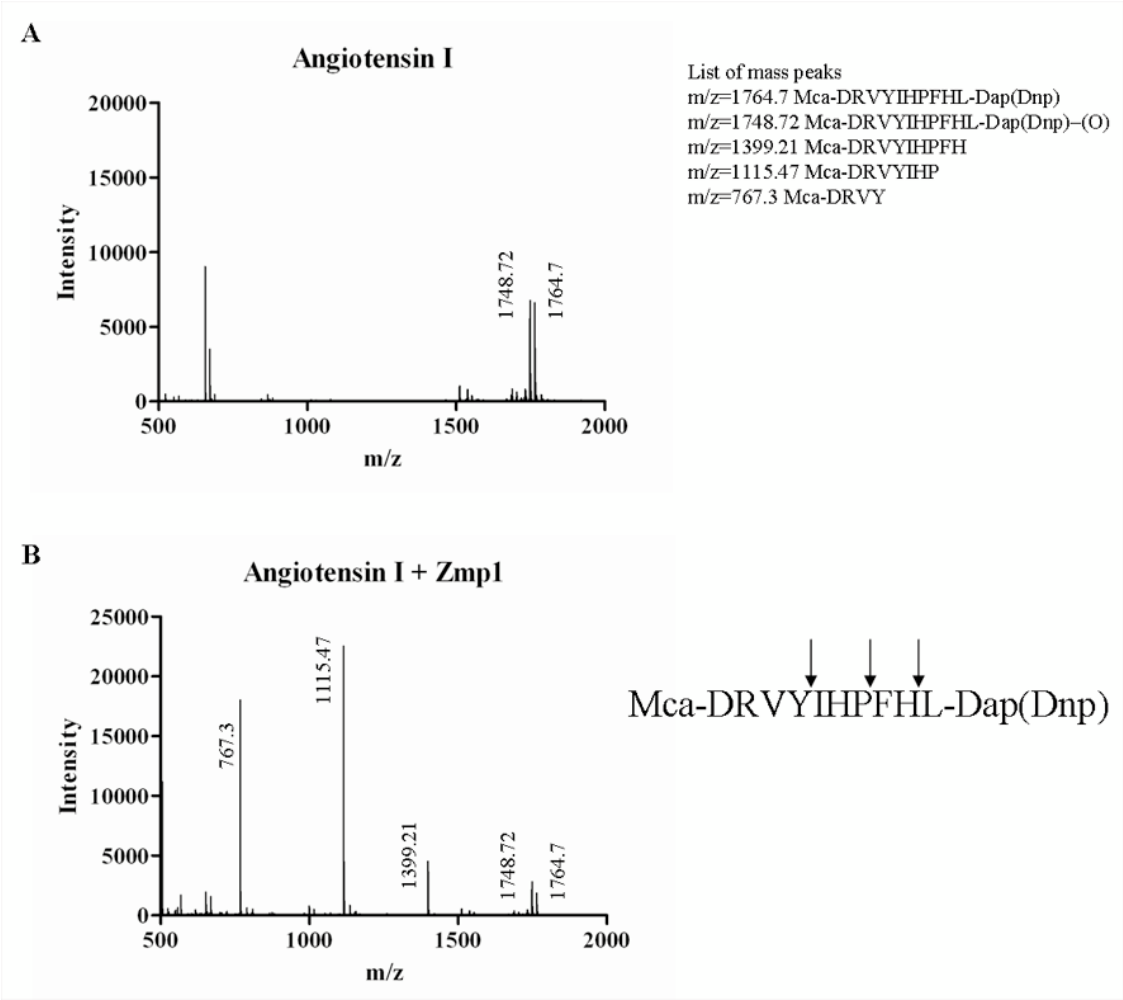


**Figure S4. Cleavage sites of apelin-13 by Zmp1**

Determination of fragments of apelin-13 by Zmp1 through LC-MS analysis. (A) MALDI-TOF spectrum of fluorogenic apelin-13. (B) MALDI-TOF spectrum of Zmp1 hydrolyzed apelin-13. The arrows above the amino acid sequence of apelin-13 indicate the cleavage sites.

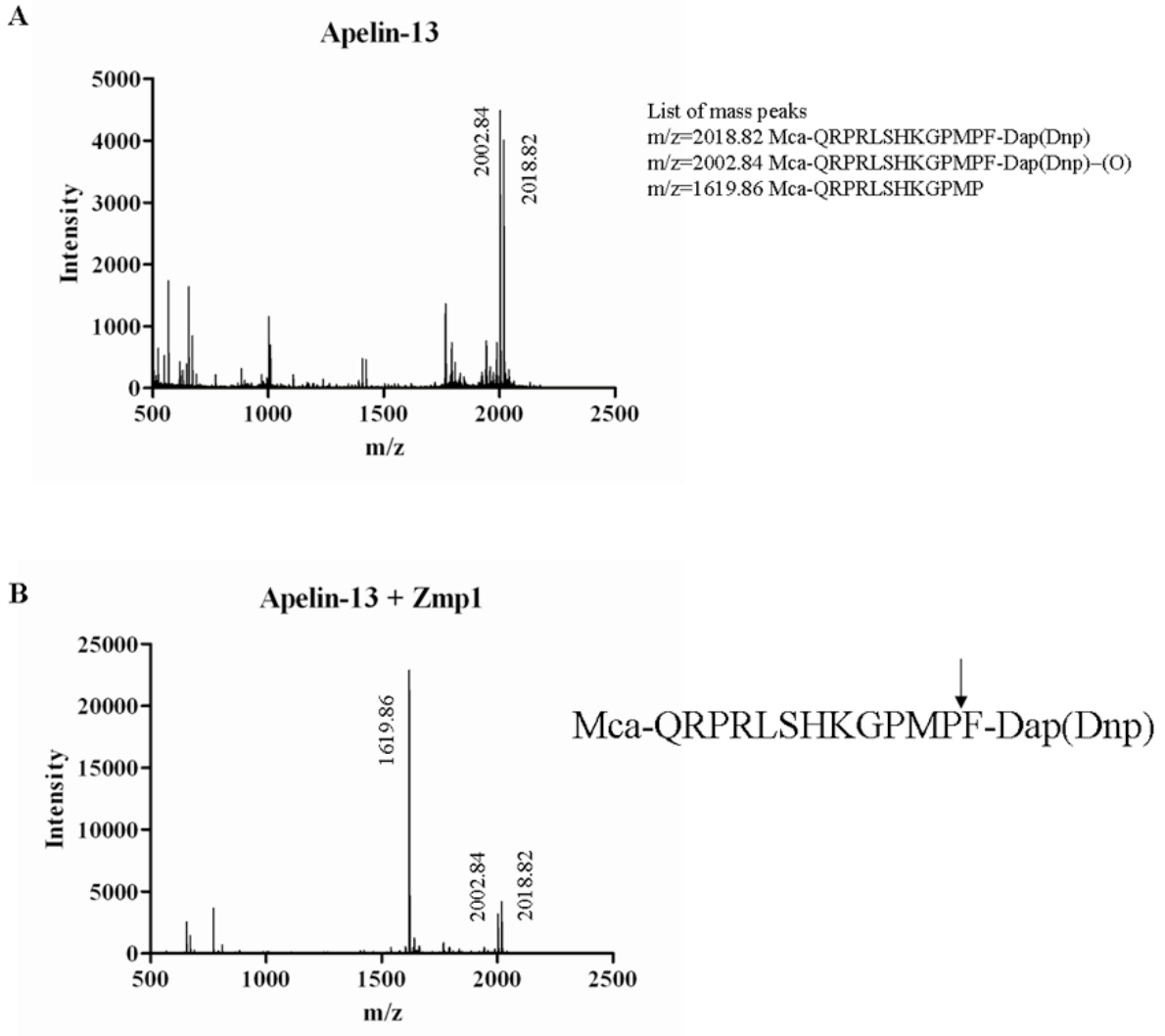


**Figure S5. Cleavage sites of dynorphin A by Zmp1**

Determination of fragments of dynorphin A by Zmp1 through LC-MS analysis. (A) MALDI-TOF spectrum of fluorogenic dynorphin A. (B) MALDI-TOF spectrum of Zmp1 hydrolyzed fluorogenic dynorphin A. The arrows above the amino acid sequence of dynorphin A indicate the cleavage sites.

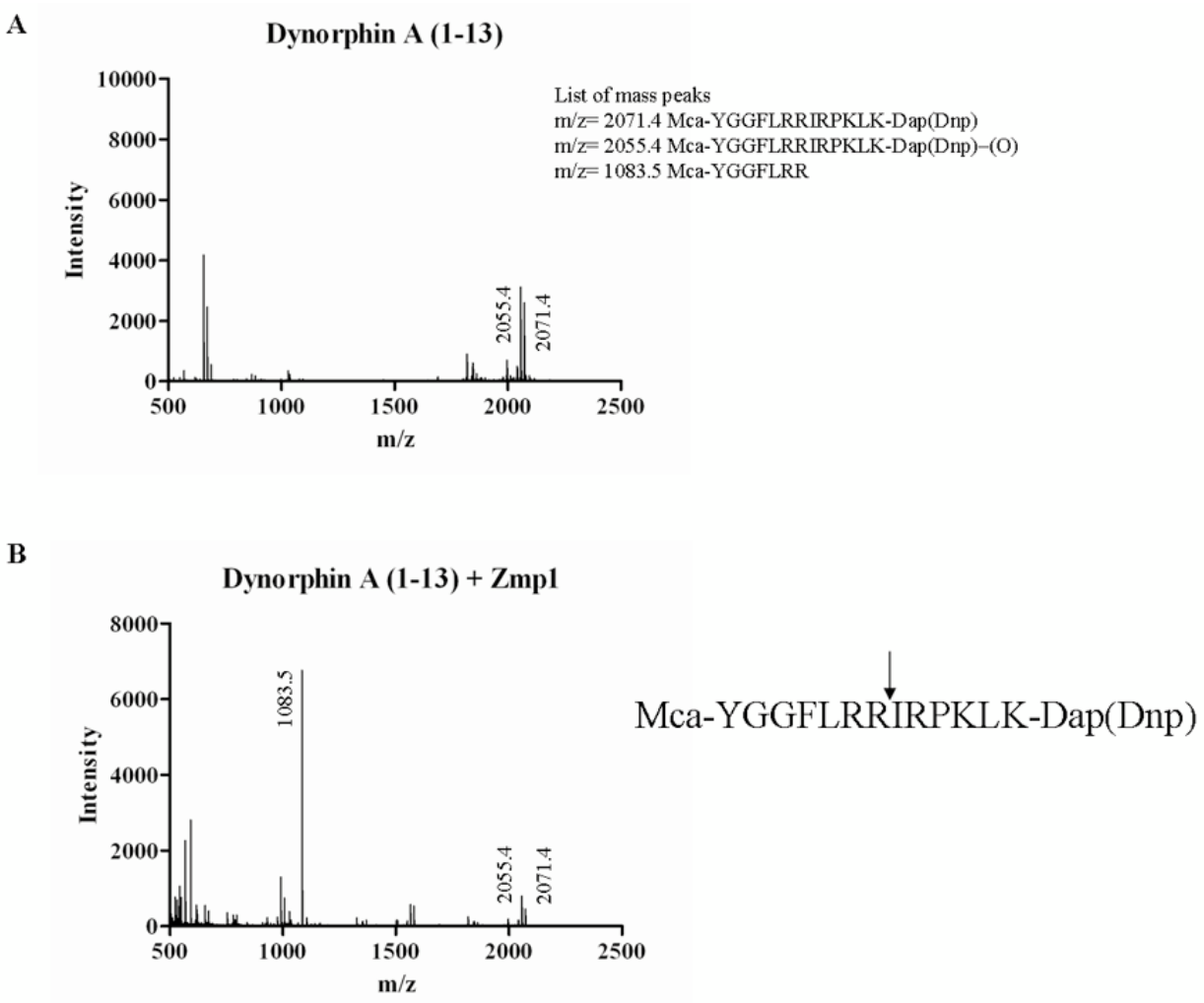
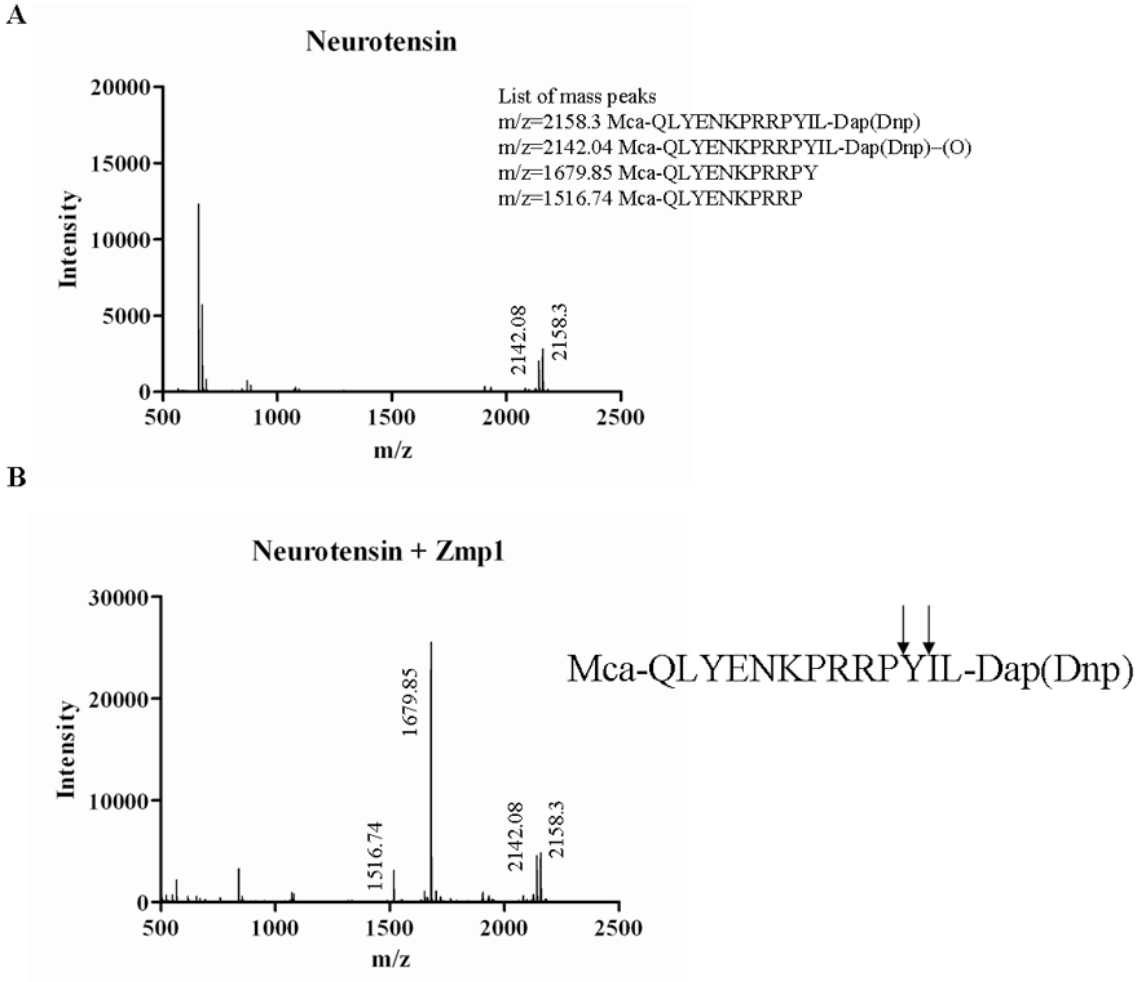


**Figure S6. Cleavage sites of neurotensin by Zmp1**

Determination of fragments of neurotensin by Zmp1 through LC-MS analysis. (A) MALDI-TOF spectrum of fluorogenic neurotensin. (B) MALDI-TOF spectrum of Zmp1 hydrolyzed fluorogenic neurotensin. The arrows above the amino acid sequence of neurotensin indicate the cleavage sites.



# PART II - Generation of three vaccine strains against tuberculosis

**Introduction**

The limited efficacy of BCG vaccine reflects the innate ability of mycobacteria to escape the host immune defence, especially the processing and presentation of antigens on the cell surface. The key for optimizing vaccination strategies against tuberculosis lies in improving access of mycobacterial antigens to presentation pathways and thus enabling high-quality T-cell priming. The genetic inactivation of BCG genes directly or indirecty involved in the suppression of antigen presentation is a promising approach for developing new vaccines, as shown by deleting the zinc metallopeptidase *zmp1* (Johansen *et al.*, 2011). The rationale of this approach is based on the hypothesis that the zinc metallopeptidase Zmp1, as player in the phagosome maturation arrest, affects the downstream antigen presentation process, particularly the MHCII-dependent pathway. Johansen and colleagues demonstrated that *zmp1* deletion increased the immunogenicity of BCG both *in vitro* and *in vivo*. Indeed, the authors showed an increased MHC class II-restricted antigen presentation in BCG Δ*zmp1* infected murine bone marrow-derived dendritic cells compared to the wild-type. As consequence of the increased antigen presentation, the enhanced BCG immunogenicity was demonstrated *in vivo*. Following immunization with BCG Δ*zmp1*, the delayed-type hypersensitivity reaction was obtained with 1/10 of the dose required for BCG wt, and the antigen-specific T-cell proliferation increased along with the cytokine secretion of splenocytes. Finally, it was proved in wild-type and SCID mice that the enhanced immunogenicity of the *zmp1* mutant is not associated with increased persistency or heightened pathology. Altogether these findings indicate an interconnection between phagosome maturation and immunogenicity. Mainly, they provide solid evidences for further developing the vaccine candidate BCG Δ*zmp1*. The mutant was generated in the substrain Pasteur, and carries an integrated kanamycin resistance cassette (Master *et al.*, 2008).

The ability of *M. tuberculosis* to arrest the phagosome maturation confers the enormous advantage to live in a protected niche isolated from the cytosol, where the

MHCI-dependent antigen presentation occurs. By impairing its ability to reside in the phagosome, S. Kaufmann created a promising vaccine strain named VPM1002 now manufactured by Vakzine Project Management GmbH. The scientific nomenclature for VPM1002 is recombinant *M. bovis* BCG *ureC::hly* (genetic background "Danish, subtype Prague"). The strain carries a hygromycin resistance marker, it is therefore resistant to this non-therapeutic antibiotic. Instead, it is sensitive to antibiotics commonly used in the treatment of mycobacterial infection, e.g. isoniazid, ethambutol and rifampicin. The recombinant BCG derives from two genetic modifications: the disruption of the urease C gene and the introduction of the lysteriolysin (Hly) gene from *Listeria monocytogenes*. The lack of Urease C contributes to maintain an acidic environment within the phagosome, essential for the activity of the pore-forming protein lysteriolysin. Mycobacterial antigens are therefore released in the cytosol, being accessible to MHCI presentation. The subsequent induction of CD8 T cells is the key for a strong and effective host immune response (Grode *et al.*, 2005), as illustrated in Figure 1.

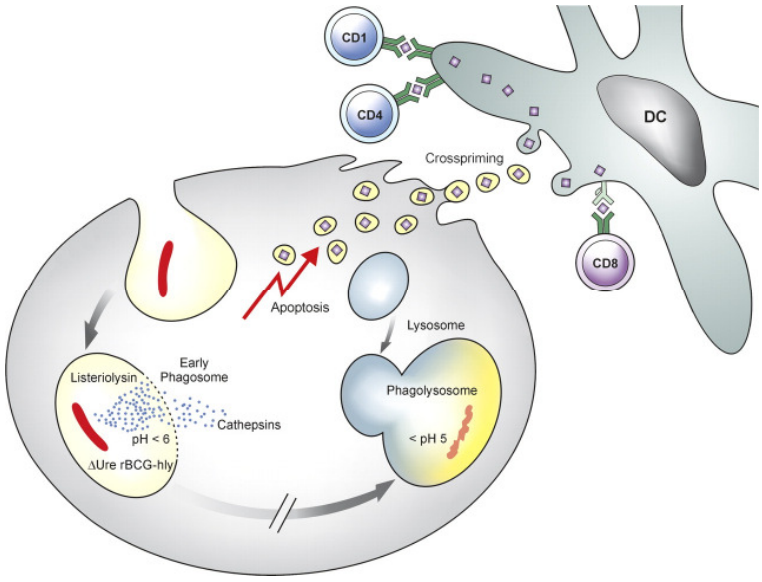



Fig.1 Mechanism of BCG *ureC::hly* vaccine strain. The recombinant strain BCG *ureC::hly* secretes lysteriolysin which provokes leakage of BCG mycobacteria and antigens into the cytoplasm. Cathepsins released in the cytosol induce apoptosis. This results in potent MHCI presentation, enhanced apoptosis, increased cross-priming and uptake by dendritic cells. The overall response involves CD8 and CD4 T cells as well as γδ T cell and CD1-restricted T cells (Kaufmann *et al.*, 2005).

The vaccine strain VPM1002 successfully completed preclinical studies, proving advanced safety profile in 10 studies comprising approximately 600 animals, from immune deficient mice to macaques (Grode *et al.*, 2005). The protective efficacy was tested in mice vaccinated with VPM1002 or BCG and subsequentially intranasally challenged using Beijing/W clinical isolate of *M. tuberculosis*. VPM1002 demonstrated improved protection against the challenge as compared to BCG (reduced CFU in the lungs by more than 100-fold at 200 days p.i.). Finally, toxicity studies performed in rodents (guinea pigs and neonatal rabbits) with high doses of VPM1002 have proved absence of adverse effect and normal weight development. The animal study findings supported the use of this vaccine in humans. VPM1002 successfully completed two Phase I clinical trials. The Phase Ia and Ib clinical trials have evaluated safety, local and systemic tolerability and immunogenicity of VPM1002 in healthy adult caucasians and healthy adult africans compared to reference control (BCG), respectively. Currently, a Phase II open label, randomized, controlled study is ongoing in South Africa to evaluate safety and immunogenicity of VPM1002 in comparison with BCG in HIV-unexposed, BCG naïve newborn infants. The preclinical and clinical safety data confirm that a single vaccination with VPM1002 up to a dose of 5x10e5 CFU is safe and well tolerated. The vaccine has shown all the characteristics of safety, tolerability and efficacy needed to replace BCG immunization in the future (Vakzine Projekt Management GmbH_VPM1002).

The BCG Δ*zmp1* and VPM1002 are two promising vaccine candidates, though the second is in a much more advanced state (Fig.2). The encouraging results obtained in mouse model with BCG Δ*zmp1* lead to further test the immunogenicity of the strain in other animal models with the perspective to enter in clinical trials. The future application of the vaccine strain in clinical trials requires the absence of resistance markers in the bacterial genome. Moreover, the *M. bovis* BCG substrain Denmark is nowadays accredited in the states of the European Union as official BCG vaccine.

The aim of this work is to create vaccine candidates with appropriate characteristics for further application in humans. Thereby, here is described the generation of the strains BCG Δ*zmp1* and BCG *ureC::hly* both unmarked, in the substrain Denmark. The combination of the properties of BCG Δ*zmp1* and BCG *ureC::hly* might further

improve BCG with respect to protective efficacy by triggering both antigen presentation pathways, MHCI and MHCII simultaneously. Therefore, we generated the combinatorial vaccine strain BCG Denmark Δ*zmp1 ureC::hly*, sensitive towards all first and second line tuberculosis drugs.

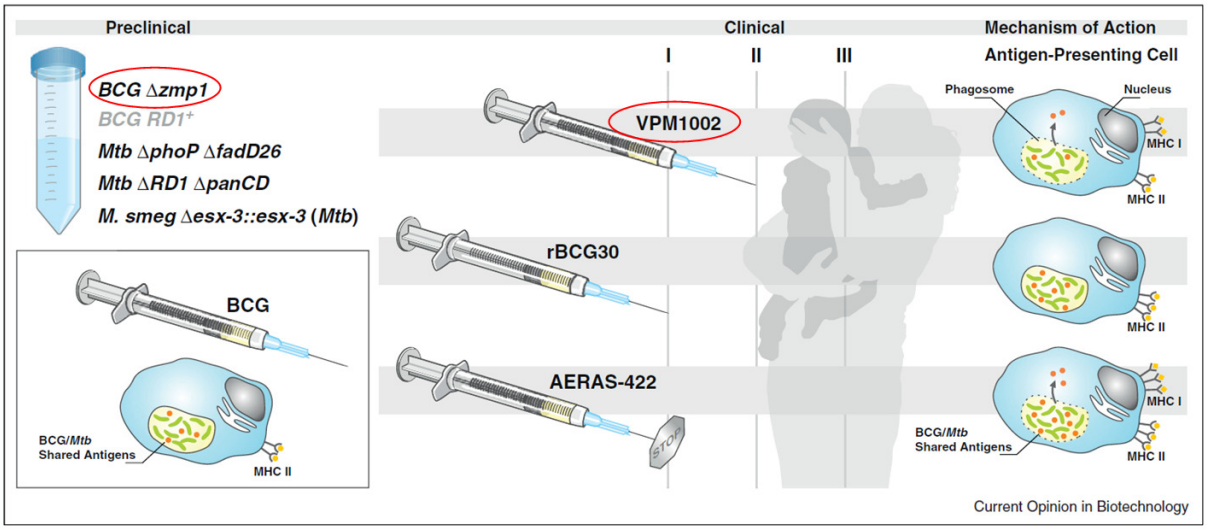



Fig.2 The current status of the vaccines BCG Δ*zmp1* and VPM1002 in the pipeline of live TB vaccines. In the boxed inset is shown the MHCII-dependent antigen presentation upon phagocytosis of BCG antigens by macrophages (or dendridic cells). BCG RD1+ will most likely not further develop because of virulence over BCG (Kaufmann and Gengenbacher, 2012).

## Experimental procedures

### Bacterial strains and growth conditions

*Escherichia coli* XL-1 blue MRF: Δ(*mcr*A)183 Δ(*mcr*CB-*hsd*SMR-*mrr*)173 *end*A1 *sup*E 44 *thi*-1 *rec*A1 *gyr*A46 *rel*A1 *lac*[F'*pro*AB *lac*I$^{q}$Z Δ M15 Tn *10*(Tet$^{r}$)]$^{C}$ (Stratagene). *E. coli* was used for initial cloning.

*M. bovis* BCG Denmark (1331) was obtained in the form of a vaccine production lot through the Statens Serum Institute, Copenhagen, Denmark.

### Bacterial growth conditions and conservation

*E. coli* was grown in LB (Luria-Bertani) broth or on LB agar. *M. bovis* BCG Denmark was grown in Middlebrook 7H9 broth or on Middlebrook 7H10 agar, supplemented with 10 % oleic acid albumin, dextrose (OADC, Difco). The strain arrived in freeze-dried (lyophilized) state and was reconstituted under laminar flow by dissolving the powder in 1 ml sterile saline (0.9 %). 10 µl of the solution were plated in 7H10 agar

plates and single colonies used for further applications. When appropriate, antibiotics were added to the media in the following concentrations: ampicillin 100 µg/ml, hygromycin 100 µg/ml for *E. coli* or hygromycin 25 µg/ml, kanamycin 50 µg/ml for *M. Bovis* BCG. *E. coli* was grown over night, *M. bovis* BCG for 3-4 weeks at 37 °C in either a plate incubator or in static cultures. *E. coli* was frozen and stored at - 80 °C in LB medium containing 10 % (v/v) glycerol. *M. bovis* BCG was frozen and stored at -80 °C in a 0.9 M NaCl solution. The vaccine strains obtained in this work were titrated and stored at -80 °C in 7H9 medium.

**Preparation of electro-competent bacteria and transformation**

*E. coli* was used for initial cloning and propagation of plasmids. An overnight culture of *E. coli* strain XL-1 blue was inoculated in 1 L of LB medium and cultivated at 37 °C until $OD_{600}$ 0.5-0.6. To prepare electrocompetent bacteria, cells were harvested by centrifugation at 5000 g for 15 min at 4 °C and washed twice with an equal volume of ice-chilled 1 mM HEPES pH 7.0. The washing procedure was repeated twice with an equal volume of 10% glycerol. Finally, bacteria were resuspended in 5-10 ml of 10% glycerol. The suspension was divided in 40 µl aliquots, frozen in liquid nitrogen and stored at -80 °C until use. For electroporation, 40 µl of competent bacteria were mixed with 2 µl of precipitated ligation reaction or 0.1 µg of supercoiled plasmid DNA (transformation control) respectively and kept on ice for 5 min. Electroporation was performed in 0.1 cm cuvettes with a single pulse ( 1.8 kV, 2 µF, 200Ω, Gene Pulser Xcell™ *BioRad*). After a recovery phase of 1 h, the bacteria were selected by plating 100 µl of electroporated bacteria on LB agar plates containing the appropriate antibiotic. Plates were incubated for 14-20 h at 37 °C.

*M. bovis* BCG Denmark was cultivated in 7H9 medium at 37 °C, and 20 ml pre-culture was used for inoculation in 200 ml 7H9 medium. The culture was grown for 25 days and one day before harvesting, BCG cells were induced with 20 ml 15% glycine. The following day cells were centrifuged at 5000 g for 20 min and resuspended in 10% glycerol. The washing and pooling was performed 4 times and finally cells were resuspended in 4 ml 10% glycerol. Aliquots of 400 µl were immediately used for electroporation. Each aliquot was mixed with 1 µg of supercoiled plasmid DNA and incubated for 5-10 min at RT. Electroporation was

performed in 2 mm cuvette with a single pulse (2.5 kV, 1000 Ω and 25 µF). After recovery phase of 24 h, the bacteria were selected by plating 2 ml of bacterial suspension on 7H10 agar plates containing the appropriate antibiotic. Plates were incubated for 3-4 weeks at 37 °C. Single colonies were picked, amplified and analyzed by PCR or Southern blotting for detection of positive clones.

**Isolation of mycobacterial genomic DNA**

*M. bovis* BCG was grown on half a plate of Middlebrook 7H10 with appropriate antibiotics. The bacteria were scraped and resuspended in 350 µL TE buffer. After a 20 min heat inactivation step at 80 °C, the bacteria solution cooled down to RT. 2 µl 20 % (v/v) Tween 80 and 10 µL lysozyme (80 mg/ml) were added to the samples, followed by an incubation step for 2 hours at 37 °C / 1000 rpm. In a second incubation step 40 µL 10 % SDS and 20 µL proteinase K (1 mg/ml) were added to the samples and incubated at 50 °C for 1 hour. For DNA extraction, 400 µl of phenol/chlorophorm/isoamylalcohol (25:24:1) was added, mixed by hand and then centrifuged at 16000 g for 20 min at 4 °C. The aqueous phase was transferred into an Eppendorf tube and the DNA precipitated by adding 8 µL of 5 M NaCl and 1 mL of ethanol. After centrifugation at 16000 g for 20 min at 4 °C, the supernatant was discarded and the pellet washed with 70 % ethanol. After drying under vacuum the DNA was resuspended in 50-100 µL of TE buffer.

**Southern Blotting**

2 µg of genomic DNA was digested with an appropriate restriction enzyme. The resulting fragments were then separated by agarose gel electrophoresis according to standard protocols. DNA was transferred to a positively charged nylon membrane (Roche) by vacuum blotting using a VacuBlot System (Biometra). Subsequently, the DNA was cross-linked to the nylon membrane by UV-irradiation using an UV-Stratalinker (Stratagene). A digoxigenin (DIG)-labelled probe was hybridized to the DNA, washed under stringent conditions, and detected by an anti-DIG antibody coupled to a horseradish peroxidase (Roche). The DNA probe was generated by PCR amplification of a 200-1000 bp DNA fragment, which was labelled with digoxigenin (DIG) according to the manufacturer's protocol.

**Targeted gene deletion in mycobacteria**

Allelic replacement technique was used to generate isogenic gene knock-outs in *M. bovis* BCG Denmark. The two suicide vectors pGEM7_sacB_hyg_aph and pMCS5_sacB_hyg, which cannot replicate in mycobacteria, were used. These vectors contain the selectable hygromycin and/or kanamycin resistance markers and *sacB*, a counterselectable marker, which kills slow-growing mycobacteria at 10% sucrose concentration in the culture medium (Pelicic *et al.*, 1996). To generate an unmarked gene deletion in *M. bovis* BCG, 1000 to 1500 bp flanking regions of the target gene were amplified, fused together and ligated into the suicide vector. The vector was transformed in BCG and selected on Middlebrook 7H10+Hyg. Bacteria that integrated the vector through homologous recombination (single cross-over event) became resistant to hygromycin and grew on the agar plates. The cross-over transformants were confirmed by PCR and Southern blot analysis. Single cross-over transformants were subjected to counterselection step. Thereby a bacterial suspension was plated on Middlebrook 7H10 supplemented of 10% sucrose. Under the sucrose pressure, only bacteria that have lost the *sacB*-containing vector through a second cross-over event will survive (Fig.3). Resulting clones were analyzed by PCR and Southern blot for allelic replacement of the targeted gene or reversion to the wild-type, respectively.

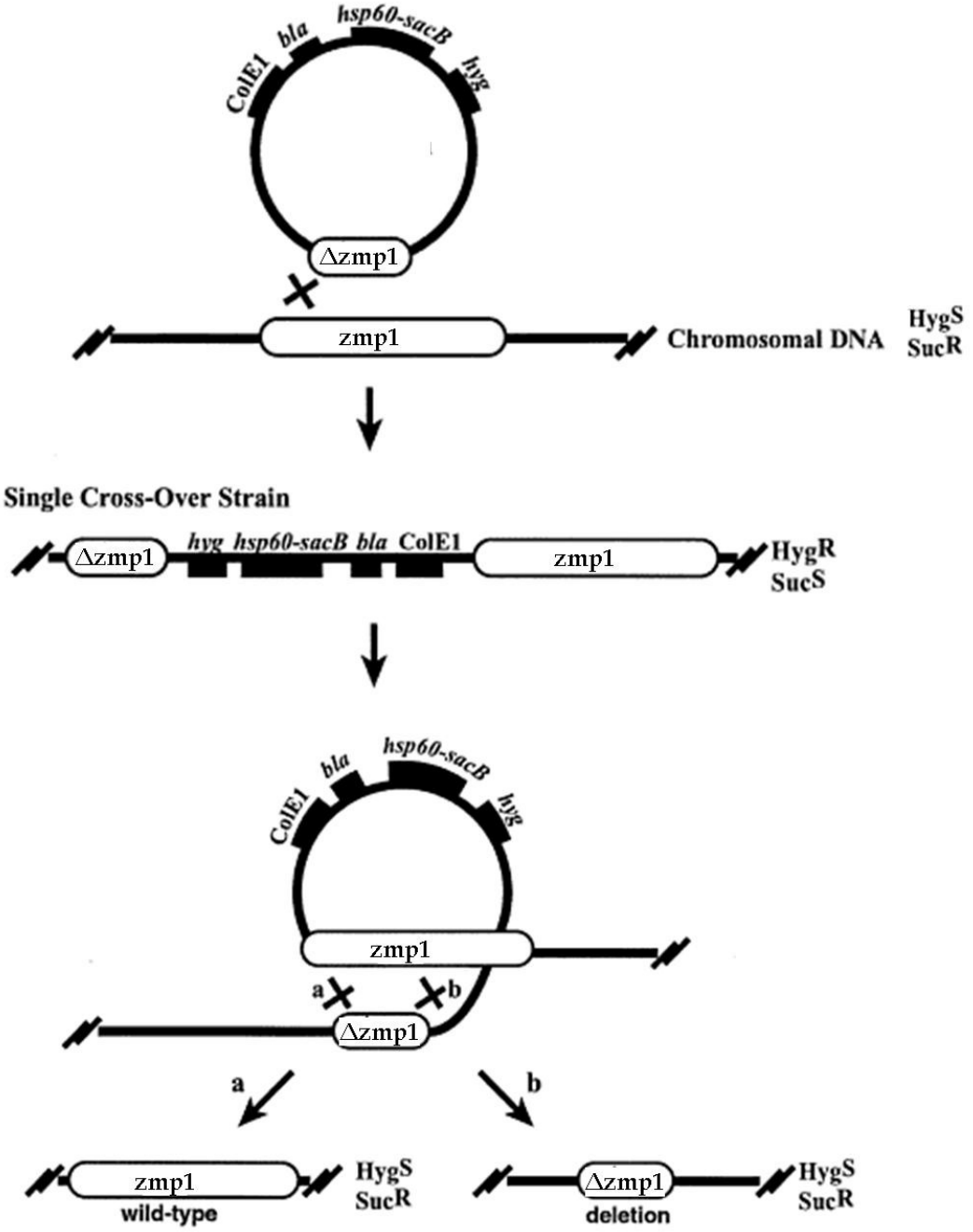



Fig.3 Schematic representation of allelic exchange via suicide counterselectable plasmid. The first event is the single homologous recombination between *zmp1* gene (as example) on the chromosome and the deleted allele present on the plasmid. This event results in the generation of a single cross-over strain. The second step involves the subsequent intrachromosomal recombination event between the wild-type *zmp1* allele and the deleted allele of *zmp1.* When the site of recombination is **a**, then the vector sequences are lost and the wild-type allele is left in the chromosome. When the site of recombination is **b**, the deleted allele remains in the chromosome and the wild-type and the vector sequences are lost (Clark, 2002).

Table 1. Primers used in this study.

a: forward primer; b: reverse primer; colored letters indicate recognition sites for restriction endonuclease enzymes

| No. | sequence | function |
|---|---|---|
| 1a | CTA GCT GAC TGA CTG AGG ATC CCG | triple translation stop insert *Nhe*I overhanged. Internal *BamH*I site |
| 2b | CTA GCG GGA TCC TCA GTC AGT CAG | triple translation stop insert *Nhe*I overhanged. Internal *BamH*I site |
| 3a | GTT GAC GGT TTG TGG TGT CAT | sequencing zmp1-stopcodons |
| 4b | TCG ACC CGT TAC TTG GTG CA | sequencing zmp1-stopcodons |
| 5a | ATG CCT GAC GGG GAG CAG | amplification of DNA probe for BCG Δ*zmp1* |
| 6b | GGA GAT CAT CAG CTT GCC G | amplification of DNA probe for BCG Δ*zmp1* |
| 7a | AGA CAA TTG GAA CAC CGC CTC GTC ACC | amplification of ureC upstream region. *Mun*I linked |
| 8b | CCG CAA TTG CCA GAG CTC CCT GCT CGA AAC CGT TGA AC | amplification of ureC upstream region. *Mun*I and *Ecl136*II linked |
| 9a | GGT GAG CTC GTC GGT GAA GGT GTC GGG A | amplification of ureC downstream region. *Ecl136*II linked |
| 10b | CCG CAA TTG AAC AGC ACT GTT TTT ATG TGT GTG A | amplification of ureC downstream region. *Mun*I linked |
| 11a | CAC CTT CCA GTG CCC CG | amplification of DNA probe for BCG *ureC::hly* |
| 12b | CGA ATG TAC TGC CCA GCT C | amplification of DNA probe for BCG *ureC::hly* |

Table 2. Plasmids used and generated in this study

| plasmid | characteristics | source |
|---|---|---|
| pGEM®-T easy | For safekeeping of PCR products, AmpR | Promega |
| pGEM7_sacB_hyg_aph | Suicide vector for homologous recombination. HygR, KanR. SacB as counterselectable marker. | (Keller *et al.*, 2008) |
| pMCS5_sacB_hyg | Suicide vector for homologous recombination. HygR. SacB as counterselectable marker. | (unpublished data) |
| pMV361_hyg | *E. coli* – mycobacteria shuttle vector, integrative in mycobacteria *attP*, *int*. KanR, HygR. Used in this study as positive control for mycobacteria transformation | (Stover *et al.*, 1991) |
| ptrpA1_rpsL_zmp | Suicide vector for homologous recombination. RpsL as counterselectable marker. It carries 4423 bp of the *zmp1* locus. Used as cloning intermediate. | (Master *et al.*, 2008) |
| ptrpA_rpsL_zmp3XSTOP | Derivate of ptrpA1-rpsL-zmp1, it carries the 3 stop codons insert within *zmp1* gene. | This study |
| pzmp_sacB_aph_hyg | Derivate of pGEM7_sacB_hyg_aph, it carries the disrupted *zmp1* gene and its flanking regions (3670 bp). | This study |
| pureC::hyg_Hly | Cosmid vector derived by pYUB854. It carries the *ureC* locus disrupted by hyg resistance marker and *hsp60-Ag85-hly* gene. | S. Kaufmann |
| pureC::Hly_sacB_hyg | Derivate of pMCS5_sacB_hyg, it carries the *ureC* locus disrupted by *hsp60-Ag85-hly* gene. | This study |

## Results

### Generation of the vaccine strain BCG Denmark Δ*zmp1* (unmarked)

Deletion of *zmp1* gene was performed by using the plasmid ptrpA1_rpsL_zmp (Master *et al.*, 2008). The plasmid carries the *M. tuberculosis* zinc metallopeptidase *zmp1* (Rv0198c; open reading frame: start 236'507; stop 234'519; 1992 bp; http://tuberculist.epfl.ch/) and its flanking regions, for a resulting fragment of 4423 bp (Fig.4). The cloning strategy is represented in Figure 5. Two adjacent *Nhe*I fragments within the *zmp1* open reading frame, comprising a total of 776 bp, were replaced by a 24-mer triple translation stop insert with an internal *BamH*I restriction site and *Nhe*I overhanging ends (Fig. 6). The 24-mer fragment was generated by annealing of oligonucleotide primers with 5'-phosphorylated ends (primer pairs: 1a and 2b). The oligonucleotides were diluted in annealing buffer (10 mM Tris-HCl, pH 8.0; 50 mM NaCl; 1mM EDTA) to a concentration 10 µM. The two oligonucleotides were mixed at equal volumes, resulting in a final concentration of 5 µM per oligonucleotide. The mixture was heated to 96° C for 5 minutes in a PCR machine and cooled down to 25° C for 8 hours. The annealing product was diluted 1:10 and used for ligation. The final product is the plasmid ptrpA_rpsL_zmp3XSTOP. The sequencing of the inserted region has confirmed the presence of the 3 stop codons. The plasmid ptrpA_rpsL_zmp3XSTOP was digested with *Not*I and the disrupted *zmp1* gene including its flanking regions (3670 bp) was cloned into *Not*I-linearized suicide vector pGEM7_sacB_aph_hyg. The resulting suicide plasmid pzmp_sacB_aph_hyg was transformed into *M. bovis* BCG Denmark, leading to generation of the knock-out strain named BCG Δ*zmp1*. The gene knock-out was confirmed by Southern blot using a 639 bp PCR fragment (primer pair: 5a and 6b) as probe (Fig.7).

The sensitivity of the strain against hygromycin, kanamycin, streptomycin and rifampicin was proved by streaking the strain on plates containing these antibiotics. The strain was then titrated and stored at -80° C for further vaccination studies in animal models.

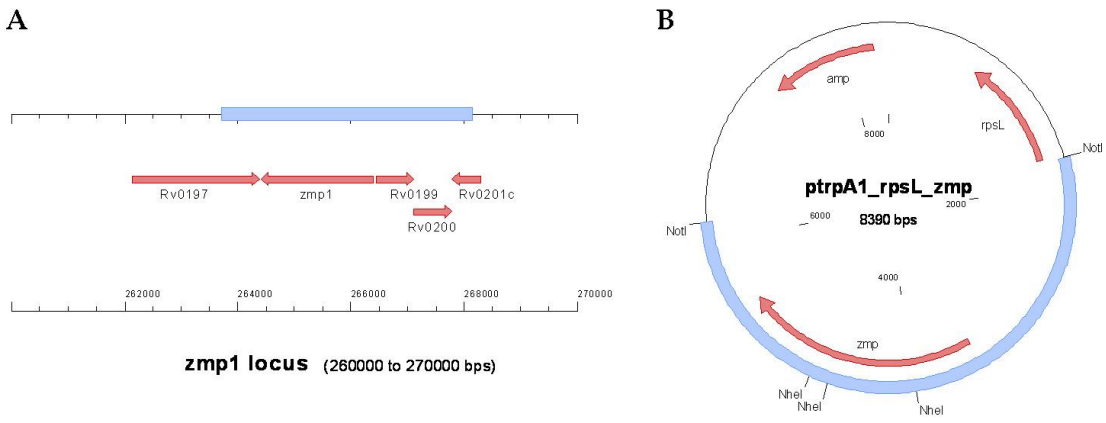



Fig. 4 A. *Zmp1* locus in *M. tuberculosis.* The blue box represents the region cloned in the vector ptrpA1_rpsL_zmp. The *M. tuberculosis zmp1* locus is equivalent in *M. bovis* BCG. Rv0197 has 99.9% identity with Mb0203, but 100% identity on amino acid level. B. Vector map of ptrpA1_rpsL_zmp.

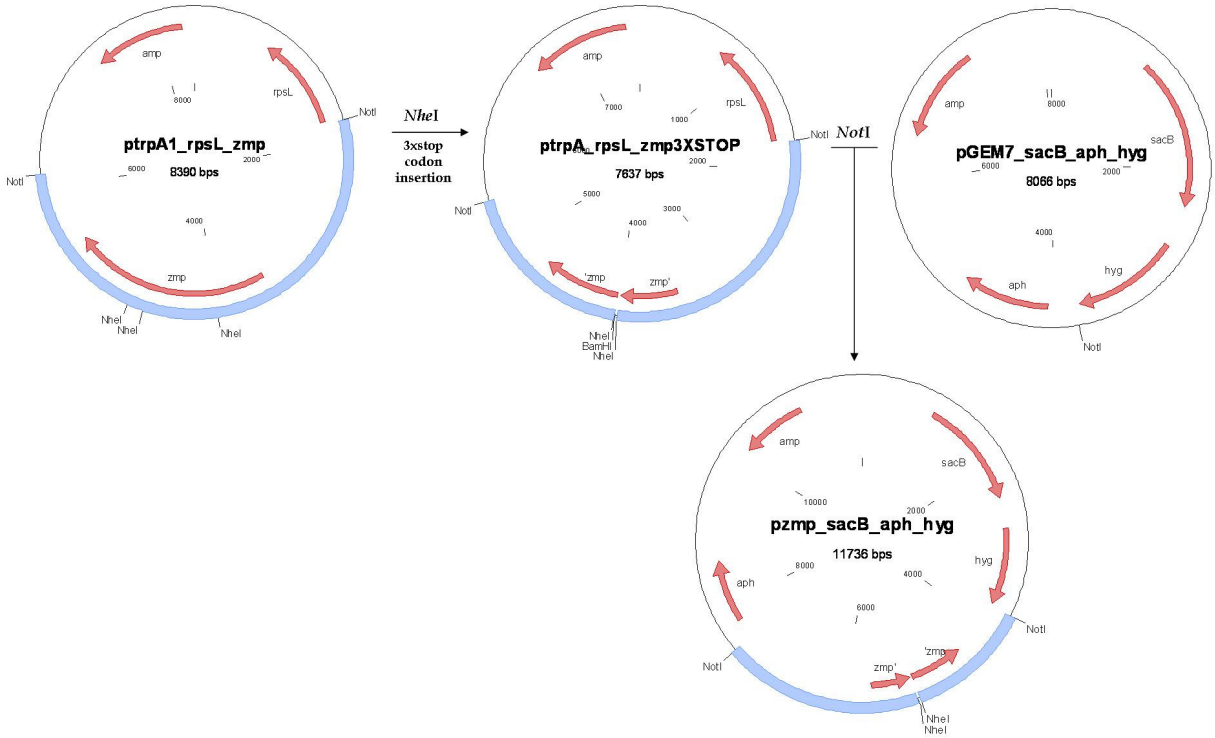



Fig.5 Cloning of the suicide vector pzmp_sacB_aph_hyg for generation of mycobacteria unmarked knock-out of *zmp1* gene.


Fig.6 24-bp triple translation stop insert with *BamH*I restriction site and *Nhe*I overhanging ends. The synthetic insert contains a stop codon for every reading frame.

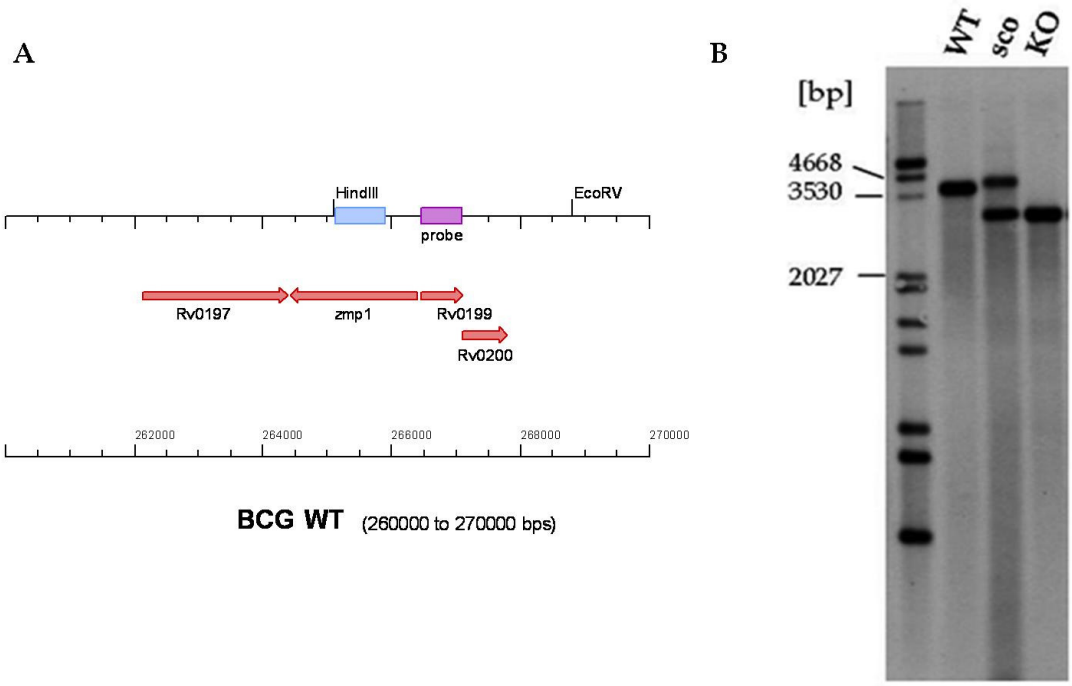



Fig. 7 Southern blot analysis confirming the deletion of *zmp1* from the genome of *M. bovis* BCG Denmark. A. Schematic drawing of the wild-type BCG *zmp1* locus. The blue box corresponds to the region which was deleted in the knock-out strain. The magenta box indicates the probe, located upstream to *zmp1* gene. The restriction endonucleases *Hind*III and *Eco*RV were used in combination to digest the genomic DNA of *M. bovis* BCG Denmark. B. The wild-type band (WT) was calculated to be 3704 bp long. The 3′ single cross-over (sco) fragments correspond to lengths of 3836 bp and 2950 bp respectively. The knock-out strain (KO) fragment corresponds to a length of 2950 bp.

**Generation of the vaccine strain BCG Denmark *ureC::hly* (unmarked)**

The vaccine strain VPM1002 was generated by transformation of the plasmid pureC::hyg_Hly in the BCG substrain Prague. This plasmid was used in this study for the generation of the unmarked vaccine strain in the substrain Denmark. The vector carries two regions named ureC P1 and ureC P2. The first region includes 658 bp upstream of the gene *ureC* (Rv1850; open reading frame: start 209'796; stop 211'530; 1734 bp; http://tuberculist.epfl.ch/) and 147 bp of the gene *ureC*, while the second region comprises 840 bp of the gene *ureC* (Fig.8). The *ureC* gene is disrupted by insertion of the hygromycin resistance marker and the gene *hly* coding the listeriolysin from *Listeria monocytogenes* fused to Antigen 85B signal peptide (for secretion of listeriolysin) and under transcriptional control of *hsp60* mycobacterial promoter.

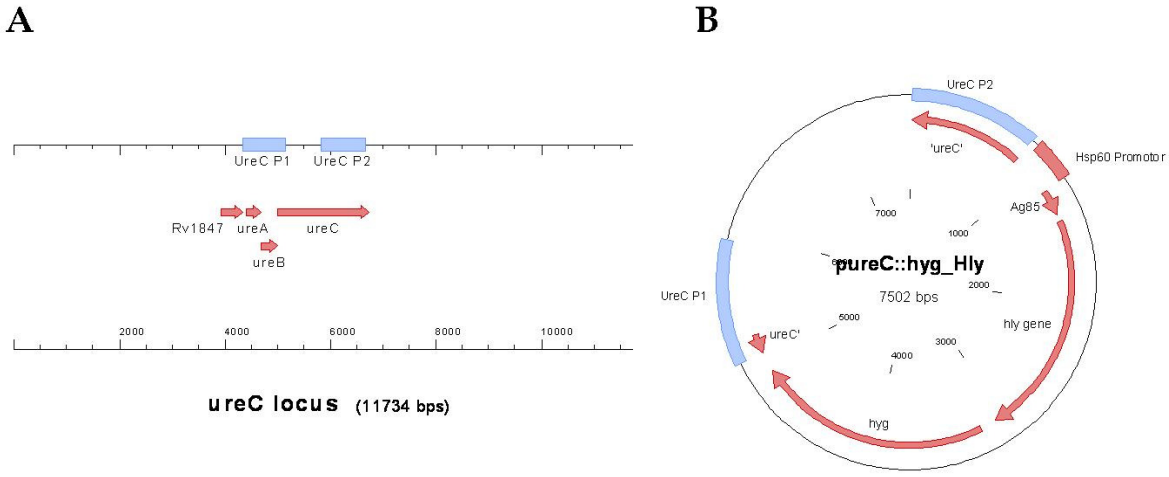



Fig. 8 A. *UreC* locus in *M. tuberculosis.* The blue boxes represent the regions ureC P1 and ureC P2 in the vector pureC::hyg_Hly. *M. tuberculosis ureC* locus is equivalent in *M. bovis* BCG. B. Vector map of pureC::hyg_Hly.

The cloning strategy is described in Fig. 9. The ureC P1 region (named upstream) was subcloned into pGEM-T easy by insertion of the amplified region (primer pairs 7a, *Mun*I linked and 8b, *Mun*I and *Ecl136*II linked), resulting in the vector UP_pGEMTeasy. The ureC P2 region along with *hsp60-Ag85-hly* gene (named downstream) was subcloned into pGEM-T easy by insertion of 3185 bp PCR product externally linked with *Ecl136*II and *Mun*I restriction endonuclease sites (primer pairs 9a and 10b). The product is the vector DOWN_pGEMTeasy. The upstream region

was excised by *Mun*I digestion and subsequentially ligated into the *Mun*I-linearized DOWN_pGEMTeasy vector, generating UP+DOWN_pGEMTeasy. The fused upstream and downstream region (total length 3990 bp) was excised by *Ecl136*II digestion and blunt ligation was performed into *EcoR*V-linearized pMCS5_hyg_sacB. The final vector, pureC::Hly_sacB_hyg, was transformed into *M. bovis* BCG Denmark, leading to generation of BCG *ureC::hly*. The gene knock-out along with the insertion of the *hly* gene was confirmed by Southern blot using a 300 bp PCR fragment (primer pair: 11a and 12b) as probe (Fig.10).

The sensitivity of the strain against hygromycin, kanamycin, streptomycin and rifampicin was proved by streaking the strain on plates containing these antibiotics. The strain was then titrated and stored at -80° C for further vaccination studies in animal models.

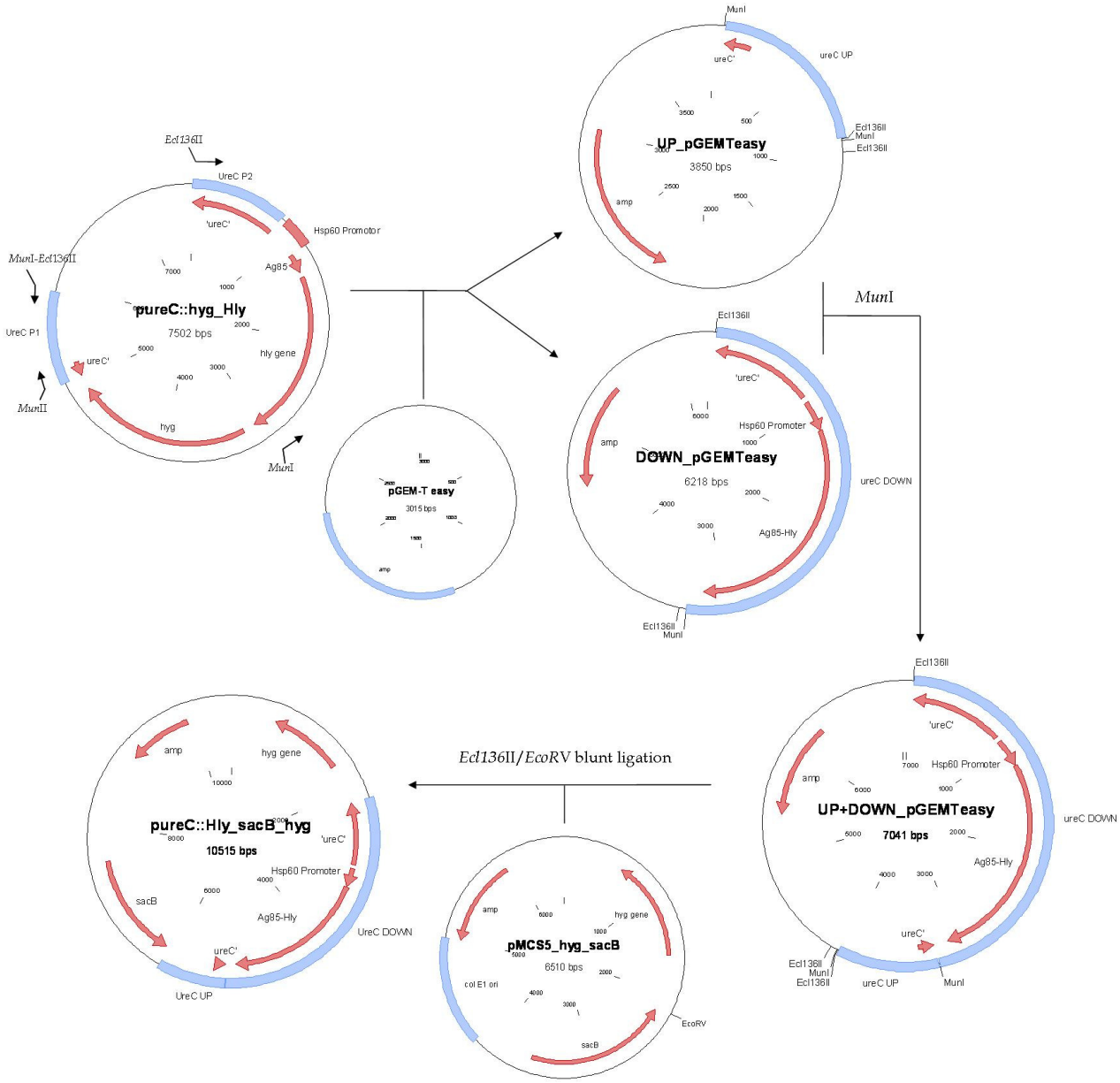



Fig.9 Cloning strategy for generation of the suicide plasmid pureC::Hly_sacB_hyg.

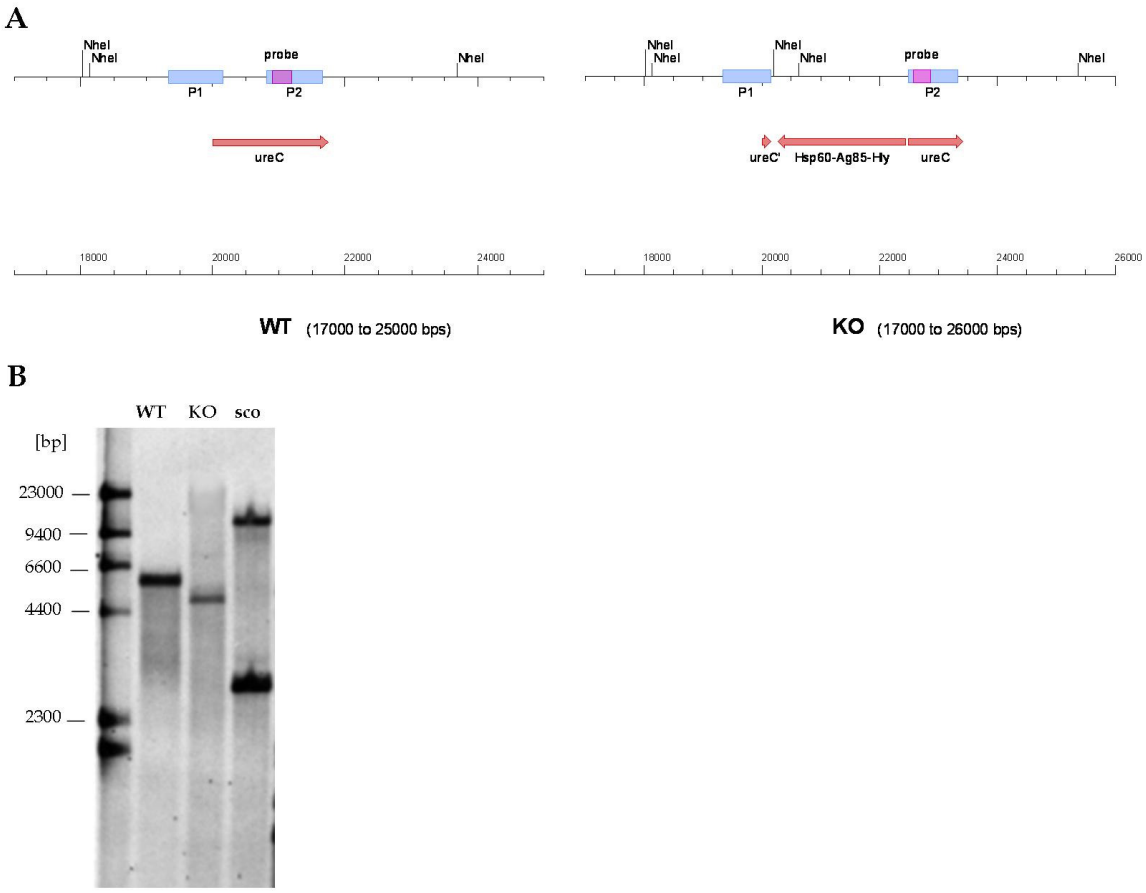



Fig. 10 Southern blot analysis confirming the deletion of *ureC* from the genome of *M. bovis* BCG Denmark and the insertion of *Ag85-hly* gene. A. Schematic drawing of the *ureC* gene in the wild-type strain (left) and of the disrupted *ureC* gene with inserted *Ag85-hly* gene in the knock-out strain (right). The blue boxes represent the ureC P1 and P2 regions used for homologous recombination. The magenta box indicates the probe, located in the ureC P2 region. The restriction endonuclease *Nhe*I was used to digest the genomic DNA of *M. bovis* BCG Denmark. B. The wild-type band (WT) was calculated to be 5546 bp long. The 5′ single cross-over (sco) fragments correspond to lengths of 2735 bp and 10842 bp respectively. The knock-out strain (KO) fragment corresponds to a length of 4741 bp.

**Generation of the vaccine strain BCG Denmark Δ*zmp1 ureC::hly* (unmarked)**

In order to combine the genetic characteristics of the two newly generated vaccine strains, the plasmid pureC::Hly_sacB_hyg (used for generating the strain BCG *ureC::hly*) was transformed into BCG Δ*zmp1.* Finally, the strain generated was named BCG Δ*zmp1 ureC::hly*. The confirmation of the genetic recombination was assessed by Southern blot applying the strategy previously used to confirm BCG *ureC::hly* (Fig.11). The strain has peculiar genetic characteristics: deletion of *zmp1* and *ureC* genes, insertion of *hly* gene and absence of antibiotic markers in the genome.

The sensitivity of the strain against hygromycin, kanamycin, streptomycin and rifampicin was proved by streaking the strain on plates containing such antibiotics. The strain was then titrated and stored at -80° C for further vaccination studies in animal models.

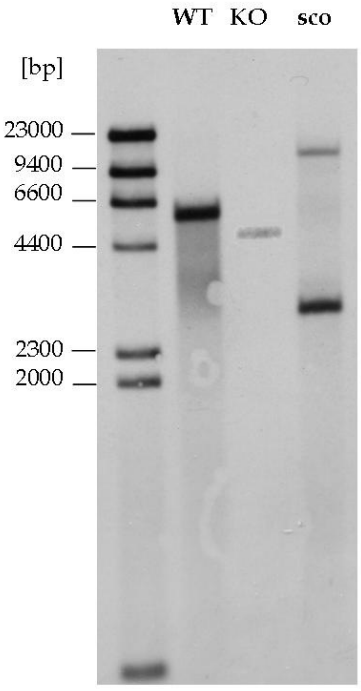



Fig. 11 Southern blot analysis confirming the deletion of *ureC* from the genome of BCG Δ*zmp1* along with the insertion of *Ag85-hly* gene. Although the background strain is BCG Δ*zmp1,* the first lane is named WT referring to *ureC* locus. The strategy used was described above for the confirmation of BCG *ureC::hly.*

**Discussion**

In the past years, it became clear in the scientific world that most of the efforts dedicated to fight tuberculosis should have been addressed in development of effective vaccines. One of the most successful vaccine candidate currently involved in clinical trials is VPM1002; its advanced safety and protection profiles drove it straight forward to Phase II trials. The strategy used to generate it proves that the escape of mycobacterial antigens in the host cytosol is a successful trick to enhance the effectiveness of vaccine. Solid evidences supporting this concept are provided by the recent work of Brosch and colleagues (Simeone *et al.*, 2012). Their fluorescence microscopy-based approach elegantly proves the translocation of *M. tuberculosis* in the host cytosol from 3 days post infection. This event was correlated to the presence of the RD1 region, and precisely to ESAT-6 which is hypothesized to have pore-forming activity. Most interestingly, BCG vaccine strain, lacking RD1 region, remains in the phagosome and by complementation with the RD1 region, it gains access to the host cytosol. The RD1-dependent phagosomal escape is the determinant of a substantially different immune response against *M. tuberculosis* and BCG. In fact, Mtb is able to induce strong CD4 T cell activation and several important cytosolic recognition pathways, such as CD8 T cells and type I interferon responses along with NLRP3 inflammasome activation (Simeone *et al.*, 2012). In contrast, BCG is an ineffectual inducer of CD8 T cells although it is still capable of stimulating CD4 T cell response (Kaufmann, 2004). It is clear from this study that a more effective vaccine strain should have the capacity to access to the cytosol in order to trigger the immune responses naturally elicited by Mtb. As unable to induce Mtb-like immune responses, the currently available BCG vaccine is uneffective.

The great potential of a strategy based on phagosomal escape was caught by S. Kaufmann, leading to the generation of a recombinant BCG provided with the pore-forming listeriolysin. The progression of the vaccine strain through clinical trials made of primary importance to generate a recombinant BCG *ureC::hly* having the characteristics needed for a future distribution in the human population. Therefore, this study provides the strain BCG Denmark *ureC::hly*, free of all antibiotic markers. The strain, generated by allelic replacement technique, was confirmed to carry the

deletion of *ureC* gene along with insertion of the *hly* gene. Although the expression of listeriolysin in the strain VPM1002 was demonstrated, as well as the leakage of mycobacterial antigens in the cytosol of infected macrophages (Grode *et al.*, 2005), these basic proofs have to be provided also for the newly generated strain. Furthermore, BCG Denmark *ureC::hly* (unmarked) has to be tested in animal models to prove the same or even better vaccine protection compared to VPM1002.

Another promising strategy for developing effective vaccines against TB is based on the zinc metallopeptidase Zmp1. Zmp1 is an important player in pathogenicity as involved in the key process of phagosome maturation arrest promoted by Mtb. Zmp1-dependent phagosome maturation arrest is conserved in BCG. Zmp1 was therefore considered a target for vaccine development. The hypothesis that the lack of the peptidase could improve the immunogenicity by enhancing the MHCII dependent-antigen presentation was validated by Joahnsen and colleagues (Johansen *et al.*, 2011). The strain BCG Δ*zmp1* was suggested as a potential vaccine candidate based on the protection and safety profile obtained in mouse model. The construction of the strain BCG Denmark Δ*zmp1* (unmarked) was performed in this study in order to replace BCG Pasteur Δ*zmp1* (marked) currently tested in guinea pigs model. The use of strains accomplishing the european requirements already in preclinical trial will accelerate future testing and will save costs and resources.

The combination of the two vaccine strains generated in this study resulted in the recombinant BCG Denmark which carries the integration of the listeriolysin gene (hly) and is devoid of Zmp1 and Urease C. These genetic modifications acting in combination are supposed to improve MHCI as well as MHCII dependent-antigen presentation, enhance apoptosis and cross-priming, potentially triggering a more potent immune response than the single modifications. It is reasonable to assume that the combined strain BCG Denmark Δ*zmp1 ureC::hly* (unmarked) will have improved protection and safety profiles, and therefore will be more successful than each of the single gene mutants.

## REFERENCES


Clark, V.L., Bavoil, P.M. (2002) *Methods in enzymology*: Academic Press.

Grode, L., Seiler, P., Baumann, S., Hess, J., Brinkmann, V., Nasser Eddine, A., Mann, P., Goosmann, C., Bandermann, S., Smith, D., Bancroft, G.J., Reyrat, J.M., van Soolingen, D., Raupach, B., and Kaufmann, S.H. (2005) Increased vaccine efficacy against tuberculosis of recombinant Mycobacterium bovis bacille Calmette-Guerin mutants that secrete listeriolysin. *J Clin Invest* **115**: 2472-2479.

Johansen, P., Fettelschoss, A., Amstutz, B., Selchow, P., Waeckerle-Men, Y., Keller, P., Deretic, V., Held, L., Kundig, T.M., Bottger, E.C., and Sander, P. (2011) Relief from Zmp1-mediated arrest of phagosome maturation is associated with facilitated presentation and enhanced immunogenicity of mycobacterial antigens. *Clin Vaccine Immunol* **18**: 907-913.

Kaufmann, S.H. (2004) New issues in tuberculosis. *Ann Rheum Dis* **63 Suppl 2**: ii50-ii56.

Kaufmann, S.H., Cole, S.T., Mizrahi, V., Rubin, E., and Nathan, C. (2005) Mycobacterium tuberculosis and the host response. *J Exp Med* **201**: 1693-1697.

Kaufmann, S.H., and Gengenbacher, M. (2012) Recombinant live vaccine candidates against tuberculosis. *Curr Opin Biotechnol*.

Keller, P.M., Bottger, E.C., and Sander, P. (2008) Tuberculosis vaccine strain Mycobacterium bovis BCG Russia is a natural recA mutant. *BMC Microbiol* **8**: 120.

Master, S.S., Rampini, S.K., Davis, A.S., Keller, C., Ehlers, S., Springer, B., Timmins, G.S., Sander, P., and Deretic, V. (2008) Mycobacterium tuberculosis prevents inflammasome activation. *Cell Host Microbe* **3**: 224-232.

Pelicic, V., Reyrat, J.M., and Gicquel, B. (1996) Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Mol Microbiol* **20**: 919-925.

Simeone, R., Bobard, A., Lippmann, J., Bitter, W., Majlessi, L., Brosch, R., and Enninga, J. (2012) Phagosomal Rupture by Mycobacterium tuberculosis Results in Toxicity and Host Cell Death. *PLoS Pathog* **8**: e1002507.

Stover, C.K., de la Cruz, V.F., Fuerst, T.R., Burlein, J.E., Benson, L.A., Bennett, L.T., Bansal, G.P., Young, J.F., Lee, M.H., Hatfull, G.F., and et al. (1991) New use of BCG for recombinant vaccines. *Nature* **351**: 456-460.

**Curriculum Vitae**

## PERSONAL DATA

| | |
|---|---|
| First Names: | Agnese |
| Name: | Petrera |
| Date of birth | March 29th, 1983 |
| Place of birth: | Rome, Italy |
| Nationality: | Italian |

## EDUCATION & RESEARCH ACTIVITIES

2008-Present: **Ph.D. thesis** at the Institute of Medical Microbiology, University of Zurich, Switzerland, under supervision of Prof. Dr. Peter Sander and Prof. Dr. Erik C. Böttger.

*Characterization of the Mycobacterium tuberculosis zinc metallopeptidase Zmp1 and generation of vaccine strains against tuberculosis*

2005-2007: **Master Degree (M.Sc.)** in Genomic Biotechnology at the Faculty of Science (MNF), University of Rome "La Sapienza", Italy, under supervision of Prof. Dr. Cristina Mazzoni.

*Analysis of the phosphoglycerate kinase (PGK1) in apoptosis and aging of the yeast Saccharomyces cerevisiae*

2002-2005: **Bachelor of Science (B.Sc.)** in Biotechnology at the Faculty of Science (MNF), University of Rome "La Sapienza", Italy.

1997-2002: **Secondary school** education (**scientific** type) at the "Talete" High School, Rome, Italy

**Conferences**

**2012**

71st Annual Assembly of the SSM (Swiss Society for Microbiology), San Gallen, Switzerland
Poster presentation:
"The *Mycobacterium tuberculosis* zinc metallopeptidase Zmp1: insights into an important player in pathogenicity"
**A. Petrera**, B. Amstutz, M. Gioia, J. Hähnlein, A. Baici, D. Ferraris, M. Rizzi, D. Sbardella, S. Marini, M. Coletta, and P. Sander

**2011**

Keystone Symposium "Tuberculosis Immunology: Cell Biology and Novel Vaccination Strategies", Vancouver, Canada
Posters presentation:

- "Characterization of the *Mycobacterium tuberculosis* zinc metallopeptidase Zmp1 and identification of its substrates"
  **A. Petrera,** J. Hähnlein, P. Selchow, D. Ferraris, M. Coletta, M. Rizzi, P. Sander and B. Amstutz
- "*Mycobacterium tuberculosis* ppm1 knock-out mutant is defective in the arrest of phagosome maturation"
  **A. Petrera,** P. Selchow, A. Tschumi, B. Amstutz, K. Walter, S. Ehlers, and P. Sander

**2010**

69th Annual Assembly of the SSM (Swiss Society for Microbiology), Zurich, Switzerland
Poster presentation:
"Characterization of the *Mycobacterium tuberculosis* zinc metallopeptidase Zmp1"
**A. Petrera**, B. Amstutz, J. Hähnlein, P. Selchow, S. Peng, D. Ferraris, M. Coletta, M. Rizzi, P. Sander

**2010**

Gordon Research Conference: Proteolytic enzymes & their inhibitors, Lucca, Italy
Poster presentation:
"Characterization of metalloproteinase Zmp1 from *Mycobacterium tuberculosis* – a protease crucial in preventing inflammasome activation and phagosome maturation"
M. Coletta**,** M. Rizzi, J. Hähnlein, B. Amstutz**, A. Petrera**, P. Selchow, P. Sander, D. Sbardella, D. Ferraris, S. Marini, V. Colizzi, M. Mattei

## Oral presentations

**2011** Biochemical and structural characterization of the *M. tuberculosis* zinc metallopeptidase Zmp1 – Progress Report, Institute of Medical Microbiology, University of Zurich

**2009** Characterization of the *M. tuberculosis* Zmp1 – Progress Report, Institute of Medical Microbiology, University of Zurich

**2008** Research proposal presentation at the Introductory Course of the PhD program in Microbiology and Immunology

## Grants

**2010** PhD Program student travel grant for participation at the Keystone Symposium "Tuberculosis Immunology: Cell Biology and Novel Vaccination Strategies" in Vancouver, Canada

## Publications

**2012** **Petrera, A**.*, Amstutz, B.*, Gioia, M.*, Hähnlein, J., Baici, A., Selchow, P., Ferraris, D.M., Rizzi, M., Sbardella, D., Marini, S., Coletta, M., Sander, P. Functional characterization of the *Mycobacterium tuberculosis* zinc metallopeptidase Zmp1 and identification of potential substrates. *Biol Chem* **393**:631-640.
* These authors contributed equally to this work

Personal contribution:

- Design of the study
- Expression and purification of the enzyme Zmp1
- Activity assays and kinetic analysis
- Data analysis of microarray peptide library
- MS analysis
- Writing of the manuscript

**2011** Ferraris, D.M., Sbardella, D., **Petrera, A**., Marini, S., Amstutz, B., Coletta, M., Sander, P., and Rizzi, M. Crystal structure of *Mycobacterium tuberculosis* zinc-dependent metalloprotease-1 (Zmp1), a metalloprotease involved in pathogenicity. *J Biol Chem* **286**: 32475-32482

Personal contribution:

- Contribution to constructing the vector expressing Zmp1 and optimization of the purification protocol
- Activity assay with fluorogenic substrate

**2009**

Mazzoni, C., Torella, M., **Petrera, A**., Palermo, V., and Falcone, C. PGK1, the gene encoding the glycolitic enzyme phosphoglycerate kinase, acts as a multicopy suppressor of apoptotic phenotypes in *S. cerevisiae*. *Yeast* **26**: 31-37

**In preparation**

Tschumi, A., Selchow, P., Laval, F., **Petrera, A**., Amstutz, B., Walter, K., Daffé, M., Ehlers, S., Sander, P.
A *Mycobacterium tuberculosis* Polyprenolmonophospho-mannose Synthase (Ppm1) mutant is defective in cell wall lipoglycans and attenuated in its immunopathogenesis

Personal contribution:

- Trafficking experiments of *M. tuberculosis* by confocal microscopy. Design of the protocol for trafficking experiment with the endosomal marker Lysotracker
- Cultivation of *M. tuberculosis* strains and mouse macrophage cell lines (RAW 264.7 and J774)